

C:Keywords: calmodulin binding; chromoprotein; FAD; flavoprotein; FMN; heme; iron; metal  
 F:533-1121/Domain: NADPH-ferrihemoprotein reductase homology <FEH>  
 F:533-671/Domain: flavodoxin homology <FLX>  
 F:194/Binding site: heme iron (Cys) (axial ligand) #status predicted

## Query Match

52.6%; Score 41; DB 1; Length 1147;

Best Local Similarity 50.0%; Pred. No. 1.1e+02;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 AAGECPALYEDPPD 14

DB 305 ADGDPPEVFEIIPD 318

## RESULT 15

S47647

nitric-oxide synthase (EC 1.14.13.39) - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 16-Jun-2000

C:Accession: S47647; JCI472

R:Geng, Y.; Almqvist, M.; Hansson, G.K.

Biochim. Biophys. Acta 1218, 421-424, 1994

A>Title: cDNA cloning and expression of inducible nitric oxide synthase from rat vascula

A:Reference number: S47647; MUID:94325351; PMID:7519448

A:Accession: S47647

A:Molecule type: mRNA

A:Residues: 1-1147 <GEN>

A:Cross-references: EMBL:X76881; NID:9439283; PIDN:CA54208.1; PID:9439284

R:Nunokawa, Y.; Ishida, N.; Tanaka, S.

Biochem. Biophys. Res. Commun. 191, 89-94, 1993

A>Title: Cloning of inducible nitric oxide synthase in rat vascular smooth muscle cells.

A:Reference number: JCI472; MUID:93191721; PMID:7680561

A:Accession: JCI472

A:Molecule type: DNA

A:Residues: 1-71, 'Y', 73-347, 'PV', 350-678, 'VP', 681-720, 'L', 722-739, 'L', 741-843, 'G', 845-10

A:Cross-references: DDBJ:D14051; NID:9286260; PIDN:BA03138.1; PID:9286261

A:Experimental source: vascular smooth muscle

C:Superfamily: nitric-oxide synthase; flavodoxin homology; NADPH-ferrihemoprotein reduct

C:Keywords: calmodulin binding; chromoprotein; FAD; flavoprotein; FMN; heme; iron; metal

F:536-1124/Domain: NADPH-ferrihemoprotein reductase homology <FEH>

F:538-674/Domain: flavodoxin homology <FLX>

F:197/Binding site: heme iron (Cys) (axial ligand) #status predicted

## Query Match

52.6%; Score 41; DB 1; Length 1147;

Best Local Similarity 50.0%; Pred. No. 1.1e+02;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 AAGECPALYEDPPD 14

DB 308 AHGDPPEVFEIIPD 321

Search completed: January 30, 2003, 14:38:54  
 Job time : 18 secs

QY 3 GEGPALYEDPP 13  
 1:111:111  
 Db 579 GGGPATEDSP 589

## RESULT 11

F88640  
 Protein F52C12.4 [imported] - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
 C:Accession: F88640  
 R:Anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology  
 A:Reference number: A75000; MUID:99069613; PMID:9851916  
 A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: F88640  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1453 <STO>  
 A:Cross-References: GB:chr\_IV; PIDN:MAC68958.1; PID:g3800950; GSPDB:GN00022; CESP:F52C12  
 C:Genetics:  
 A:Gene: F52C12.4  
 A:Map position: 4

Query Match 53.8%; Score 42; DB 2; Length 1453;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LYEDPPD 14  
 1111111  
 Db 517 LYEDPPD 523

## RESULT 12

A84567  
 hypothetical protein At2g18660 [imported] - *Arabidopsis thaliana*  
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: A84567  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
 Neuss, D.; Niernman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: A84567  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-126 <STO>  
 A:Cross-References: GB:AE002093; NID:g4185132; PIDN:AAD08935.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g18660  
 A:Map position: 2

Query Match 52.6%; Score 41; DB 2; Length 126;  
 Best Local Similarity 58.3%; Pred. No. 9.5;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 AGEPPALYEDPP 13  
 1:111:111  
 Db 22 AAGKAYIDPP 33

## RESULT 13

AF0617  
 probable DNA methylase STY1014 [imported] - *Salmonella enterica* subsp. *enterica* serovar  
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh  
 A:Note: this species has also been called *Salmonella typhi*  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 02-Aug-2002  
 C:Accession: AF0617  
 R:Parhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr  
 S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,  
 A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* se  
 A:Reference number: AB0502; PMID:11677608  
 A:Accession: AF0617  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-293 <PAR>  
 A:Cross-References: GB:AL513382; PIDN:CAD05408.1; PID:g16502169; GSPDB:GN00176  
 C:Genetics:  
 A:Gene: STY1014  
 C:Superfamily: mbra protein

Query Match 52.6%; Score 41; DB 2; Length 293;  
 Best Local Similarity 58.3%; Pred. No. 25;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 AGEPPALYEDPP 13  
 111:111:111  
 Db 193 AGDGVYCDPP 204

## RESULT 14

AA3271  
 nitric-oxide synthase (EC 1.14.13.39), calmodulin-independent - mouse  
 C:Species: *Mus musculus* (house mouse)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 03-Mar-2000  
 C:Accession: AA3271; A42166; JN0458; A46186  
 R:Xie, Q.; Cho, H.J.; Calaycay, J.; Mumford, R.A.; Swiderek, K.M.; Lee, T.D.; Ding, A  
 Science 256, 225-228, 1992  
 A:Title: Cloning and characterization of inducible nitric oxide synthase from mouse m  
 A:Reference number: AA3271; MUID:92229444; PMID:1373522  
 A:Accession: AA3271  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1144 <XIE>

A:Cross-References: GB:M87039; NID:g198406; PIDN:AAA39315.1; PID:g198407  
 R:Lyons, C.R.; Orloff, G.J.; Cunningham, J.M.  
 J. Biol. Chem. 267, 6370-6374, 1992  
 A:Title: Molecular cloning and functional expression of an inducible nitric oxide syn

A:Reference number: A42166; MUID:92210618; PMID:1372907  
 A:Accession: A42166  
 A:Status: preliminary

A:Molecule type: mRNA  
 A:Residues: 1-1144 <LYO>  
 A:Cross-References: GB:M84373; NID:g200095; PIDN:AAA39834.1; PID:g200096

R:Wood, E.R.; Berger, Jr., H.; Sherman, P.A.; Lapetina, E.G.  
 Biochem. Biophys. Res. Commun. 191, 767-774, 1993  
 A:Title: Hepatocytes and macrophages express an identical cytokine inducible nitric o

A:Reference number: JN0457; MUID:93221515; PMID:7682072  
 A:Accession: JN0458  
 A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA  
 A:Residues: 1-278 'E', 280-682, 'H', 684-937, 939-1144 <MO>  
 A:Experimental source: liver

R:Lowenstein, C.J.; Glatt, C.S.; Bredt, D.S.; Snyder, S.H.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 6711-6715, 1992  
 A:Title: Cloned and expressed macrophage nitric oxide synthase contrasts with the bra

A:Reference number: A46186; MUID:92357701; PMID:1379716  
 A:Accession: A46186  
 A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA  
 A:Residues: 1-190, 'V', 192-765, 'P', 767-843, 'G', 845-1144 <LOW>  
 A:Cross-References: GB:M92649; NID:g200109

A:Experimental source: BALB/c, RAW 264.7 cells, macrophage  
 A:Note: sequence extracted from NCBI backbone (NCBIP:113541)  
 C:Genetics:  
 A:Gene: NOS  
 C:Function:  
 A:Description: catalyzes the oxidation of an L-arginine guanidino nitrogen and of NAD  
 C:Superfamily: nitric-oxide synthase; flavodoxin homology; NADPH-ferrihemoprotein red

## RESULT 8

Hypothetical protein DKFZp434L0850.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 18-Feb-2000

C:Accession: T42699; T34559

R:Postik, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, November 1999

A:Reference number: Z22232

A:Accession: T42699

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-303 <AA>

A:Cross-references: EMBL:AL133042

A:Experimental source: adult testis; clone DKFZp434L0850

R:Postik, A.; Wellenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, October 1999

A:Reference number: Z21540

A:Accession: T34559

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-303 <PO>

A:Cross-references: EMBL:AL122038

A:Experimental source: adult testis; clone DKFZp434L0850

C:Genetics:

A:Note: DKFZp434L0850.1

## Query Match

Best Local Similarity 53.8%; Score 42; DB 2; Length 303;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 AAGECPALYEDP 12  
:|||||:|  
DB 84 SAGEGPVIVHP 95

## RESULT 9

Nitric-oxide synthase (EC 1.14.13.39), inducible - human

C:Species: Homo sapiens (man)

C>Date: 16-Feb-1996 #sequence\_revision 16-Feb-1996 #text\_change 20-Jun-2000

C:Accession: A49676; JX0345; G01947; J38933; S47566; A47475

R:Charles, I.G.; Palmer, R.M.; Hickery, M.S.; Bayliss, M.T.; Chubb, A.P.; Hall, V.S.; Mc

Proc. Natl. Acad. Sci. U.S.A. 90, 11419-11423, 1993

A:Title: Cloning, characterization, and expression of a cDNA encoding an inducible nitric

A:Reference number: A49676; MUID:94068614; PMID:7504305

A:Accession: A49676

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1153 <RES>

A:Cross-references: EMBL:X73029; NID:9441452; PIDN:CAA51512.1; PID:9441453

R:Hokari, A.; Zeniya, M.; Esumi, H.

J. Biochem. 116, 575-581, 1994

A:Title: Cloning and functional expression of human inducible nitric oxide synthase (NOS

A:Reference number: JX0345; MUID:95155267; PMID:7531687

A:Accession: JX0345

A:Molecule type: mRNA

A:Residues: 1-607, 'L', 609-1153 <HOK>

A:Cross-references: DBJ:D26525; NID:9559326; PIDN:BA05531.1; PID:g1228940

A:Experimental source: glioblastoma cell line A-172

R:Park, C.; Park, R.; Krishna, G.

submitted to the EMBL Data Library, July 1995

A:Reference number: G08912

A:Accession: G01947

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-607, 'L', 609-1153 <PAR>

A:Cross-references: EMBL:U31511; NID:9591320; PIDN:AAB49041.1; PID:9591321

R:Guo, F.H.; De Raeye, H.R.; Rice, T.W.; Stuehr, D.J.; Thummler, F.B.; Erratum, S.C.

Proc. Natl. Acad. Sci. U.S.A. 92, 7809-7813, 1995

A:Title: Continuous nitric oxide synthase by inducible nitric oxide synthase in normal

A:Reference number: J38933; MUID:95372368; PMID:7544004

A:Accession: J38933

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-675, 'I', 677-932, 'G', 934-965, 'A', 967-1153 <RES>

A:Cross-references: EMBL:U20141; NID:9687680; PIDN:AAB60366.1; PID:9687681

R:Maier, R.; Bilbe, G.; Rediske, J.; Lotz, M.

Biochim. Biophys. Acta 1208, 145-150, 1994

A:Title: Inducible nitric oxide synthase from human articular chondrocytes: cDNA clone

A:Reference number: S47566; MUID:94368816; PMID:7522054

A:Accession: S47566

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-22, 'G', 24-153, 'L', 155-176, 'V', 178-799, 'A', 801-912, 'P', 914-1153 <MAI>

A:Cross-references: EMBL:U05810; NID:94552487; PIDN:AA56666.1; PID:94552488

R:Geller, D.A.; Lowenstein, C.J.; Shapiro, R.A.; Nussler, A.K.; Di Silvio, M.; Wang,

Proc. Natl. Acad. Sci. U.S.A. 90, 3491-3495, 1993

A:Title: Molecular cloning and expression of inducible nitric oxide synthase from hum

A:Reference number: A47475; MUID:9334523; PMID:7682706

A:Accession: A47475

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-422, 'I', 424-804, 'D', 806-830, 'SP', 833-932, 'G', 934-965, 'A', 967-986, 'V', 98

A:Cross-references: GB:L09210; NID:9292241; PIDN:AA59171.1; PID:9292242

A:Experimental source: hepatocytes

A:Note: sequence extracted from NCBI backbone (NCBI:129733)

C:Genetics:

A:Gene: GDB:NOS2A; NOS2; INOS

A:Cross-references: GDB:139215; OMIM:163730

A:Map position: 17cen-17q11.2

C:Function:

A:Description: catalyzes the oxidation of an L-arginine guanidino nitrogen and of NAD

C:Superfamily: nitric-oxide synthase; flavodoxin homology; NADPH-ferritinoprotein red

C:Keywords: calmodulin binding; chromoprotein; FAD; Flavoprotein; FMN; heme; iron; me

F:509-529/Region: calmodulin binding #status predicted

F:539-1127/Domain: NADPH-ferritinoprotein reductase homology <FLX>

F:541-677/Domain: flavodoxin homology <FLX>

F:623-654/Region: FMN binding #status predicted

F:765-778/Region: FAD-pyrophosphate binding #status predicted

F:903-913/Region: FAD-isoloxazine binding #status predicted

F:978-996/Region: NADP-ribose binding #status predicted

F:1076-1091/Region: NADP-adenine binding #status predicted

F:200/Binding site: heme iron (Cys) (axial ligand) #status predicted

## Query Match

Best Local Similarity 53.8%; Score 42; DB 2; Length 1153;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 AAGECPALYEDPPD 14  
| | | | | | | |  
DB 311 ANGRDPELEFIRPPD 324

## RESULT 10

hairless protein - mouse

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999

C:Accession: I48378

R:Cachon-Gonzalez, M.B.; Fennel, S.; Coffin, J.M.; Moran, C.; Best, S.; Stoye, J.P.

Proc. Natl. Acad. Sci. U.S.A. 91, 7717-7721, 1994

A:Title: Structure and expression of the hairless gene of mice.

A:Reference number: I48378; MUID:9429587; PMID:8052649

A:Accession: I48378

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1182 <RES>

A:Cross-references: EMBL:X2675; NID:9531706; PIDN:CAA83587.1; PID:9531707

## Query Match

Best Local Similarity 53.8%; Score 42; DB 2; Length 1182;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

C;Keywords: carbon-oxygen lyase; DNA repair; endonuclease; nucleus  
F;2-317/Product: DNA repair enzyme APex nuclease #status experimental <MAT>

## Query Match

Best Local Similarity 94.9%; Score 74; DB 2; Length 317;  
Pred. No. 5.9e-05;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAGEGPAHYEDPPD 14

||||| |||||||

DB 36 AAGEGPAHYEDPPD 49

## RESULT 4

S26830

DNA-(apurinic or apyrimidinic site) lyase (EC 4.2.99.18) - bovine  
N;Alternate names: apurinic/apyrimidinic endonuclease I; BAP-1 protein

C;Species: Bos primigenius taurus (cattle)

C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 18-Feb-2000

C;Accession: S26830; S30761

R;Robson, C.N.; Milne, A.M.; Pappin, D.J.C.; Hickson, I.D.

Nucleic Acids Res. 19, 1087-1092, 1991

A;Title: Isolation of cDNA clones encoding an enzyme from bovine cells that repairs oxid

A;Reference number: S26830; MUID:91212189; PMID:1708495

A;Accession: S26830

A;Molecule type: mRNA

A;Residues: 1318 <ROB>

A;Cross-references: EMBL:X56685; NID:g117; PIDN:CAA40014.1; PID:g118

A;Note: part of this sequence, including the amino end of the mature protein, was confir

R;Hennner, W.D.; Kiker, N.P.; Jorgensen, T.J.; Munch, J.N.

Nucleic Acids Res. 15, 5529-5544, 1987

A;Title: Purification and amino-terminal amino acid sequence of an apurinic/apyrimidinic

A;Reference number: S30761; MUID:87289027; PMID:2441359

A;Accession: S30761

A;Molecule type: protein

A;Residues: 2-20, 'Lp', 23 <HEN>

C;Superfamily: exodeoxyribonuclease III

C;Keywords: carbon-oxygen lyase; DNA repair; endonuclease; nucleus

F;2-318/Product: DNA-(apurinic or apyrimidinic site) lyase #status predicted <MAT>

## Query Match

Best Local Similarity 79.5%; Score 62; DB 2; Length 318;  
Pred. No. 0.0067;

Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAGEGPAHYEDPPD 14

||||| |||||||

DB 37 AAGEGPAHYEDPPD 50

## RESULT 5

U13600

C;Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 19-Apr-2002

C;Accession: A13600

R;DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

J.; Mazur, M.; Goldsman, E.; Selkov, P.H.; Hagius, S.; O'Callaghan, D.; Letess

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A;Reference number: AD3252; PMID:11756688

A;Accession: A13600

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-328 <KUR>

A;Cross-references: GB:AE008918; PIDN:AAU53972.1; PID:g17984919; GSPDB:GN00191

A;Experimental source: strain 16M

C;Genetics:

A;Gene: BMEI10730

A;Map position: II

C;Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

## Query Match

Best Local Similarity 56.4%; Score 44; DB 2; Length 328;  
Pred. No. 8.5;

DB 37 AAGEGPAHYEDPPD 50

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAGEGPAHYEDP 12

||||| |||||

DB 285 AGDPPALYADP 295

## RESULT 6

C72293

hypothetical protein - Thermotoga maritima (strain MSB8)

C;Species: Thermotoga maritima

C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 28-Jul-2000

C;Accession: C72293

R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hlc

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,

C.M.

Nature 399, 323-329, 1999

A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome

A;Reference number: A72200; MUID:99287316; PMID:10360571

A;Accession: C72293

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-612 <ARN>

A;Cross-references: GB:AE001770; GB:AE000512; NID:g4981658; PIDN:AA036202.1; PID:g498

A;Experimental source: strain MSB8

C;Genetics:

A;Gene: TM1126

C;Superfamily: Thermotoga maritima hypothetical protein TM1126

## Query Match

Best Local Similarity 55.8%; Score 43.5; DB 2; Length 612;  
Pred. No. 21;

Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 2 AAGEGPAHYEDPP 13

||||| |||||

DB 224 AAGEGPR-YENPP 234

## RESULT 7

S20473

fatty-acid synthase (EC 2.3.1.85) - Brevibacterium ammoniagenes

C;Species: Brevibacterium ammoniagenes

C;Date: 22-Nov-1993 #sequence\_revision 12-Apr-1996 #text\_change 05-May-2000

C;Accession: S20473; S28645

R;Meurer, G.; Biermann, G.; Schuetz, A.; Harth, S.; Schweizer, E.

Mol. Gen. Genet. 232, 106-116, 1992

A;Title: Molecular structure of the multifunctional fatty acid synthetase gene of Bre

he two yeast genes FAS1 and FAS2.

A;Reference number: S20473; MUID:92204122; PMID:1552898

A;Accession: S20473

A;Molecule type: DNA

A;Residues: 1-3104 <MEU>

A;Cross-references: EMBL:X64795

R;Schweizer, E.

submitted to the EMBL Data Library, December 1992

A;Reference number: S28645

A;Accession: S28645

A;Molecule type: DNA

A;Residues: 1-1430, 'V', 1432-3104 <SCH>

A;Cross-references: EMBL:X64795; NID:g39319; PIDN:CAA46024.1; PID:g580746

C;Genetics:

A;Gene: FAS

A;Start codon: GTG

C;Superfamily: Brevibacterium ammoniagenes fatty-acid synthase

C;Keywords: acyltransferase; coenzyme A

## Query Match

Best Local Similarity 55.1%; Score 43; DB 2; Length 3104;  
Pred. No. 1.0e+02;

Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAGEGPAHYEDPPD 14

||||| |||||

DB 2960 A1SGGPAHYERPD 2973

A:Residues: 1-50, 'H', 52-318 <SEK>  
 A:Cross-references: DDBJ:DJ3370; NID:9219473; PIDN:BA02633.1; PID:9219474  
 R:Akiyama, K.; Seki, S.; Oshida, T.; Yoshida, M.C.  
 Biochim. Biophys. Acta 1219, 15-25, 1994  
 A:Title: Structure, promoter analysis and chromosomal assignment of the human APEX gene.  
 A:Reference number: 547521; MUID:94368844; PMID:8086453  
 A:Accession: 547521  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-147, 'E', 149-318 <AKT>  
 A:Cross-references: DDBJ:DJ3370; NID:9219473; PIDN:BA02633.1; PID:9219474  
 R:Harlison, L.; Asclone, G.; Menninger, J.C.; Ward, D.C.; Dimple, B.  
 Hum. Mol. Genet. 1, 677-680, 1992  
 A:Title: Human apurinic endonuclease gene (APE): structure and genomic mapping (chromosome 10p11.23)  
 A:Reference number: 139472; MUID:93258307; PMID:1284593  
 A:Accession: 139472  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-146 <HAR>  
 A:Cross-references: GB:M9703; NID:9178748; PIDN:AAA58373.1; PID:9553182  
 R:Robson, C.N.; Hickson, I.D.  
 Nucleic Acids Res. 19, 5519-5523, 1991  
 A:Title: Isolation of cDNA clones encoding a human apurinic/aprimidinic endonuclease the  
 A:Reference number: 534422; MUID:92051291; PMID:1719477  
 A:Accession: 534422  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-236, 'A', 238-318 <RO2>  
 A:Cross-references: EMBL:X5764; NID:932023; PIDN:CAA2437.1; PID:932024  
 R:Dimple, B.; Herman, T.; Chen, D.S.  
 Proc. Natl. Acad. Sci. U.S.A. 88, 11450-11454, 1991  
 A:Title: Cloning and expression of APE, the cDNA encoding the major human apurinic endonuclease  
 A:Reference number: A41631; MUID:92107968; PMID:1722334  
 A:Accession: A41631  
 A:Molecule type: mRNA  
 A:Residues: 1-56, 'A', 58-305, 'A', 307-318 <DEM1>  
 A:Cross-references: GB:M80261; NID:9178742; PIDN:AAA58371.1; PID:9178743  
 A:Accession: B41631  
 A:Molecule type: protein  
 A:Residues: 2-21, 'P', 23-26 <DEM2>  
 R:Barzilay, G.; Walker, L.J.; Robson, C.N.; Hickson, I.D.  
 Nucleic Acids Res. 23, 1544-1550, 1995  
 A:Title: Site-directed mutagenesis of the human DNA repair enzyme HAP1: identification of  
 A:Reference number: S55547; MUID:95303638; PMID:7784208  
 A:Accession: S55547  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 62-318 <BAR>  
 C:Genetics:  
 A:Gene: GDB:APX; ref-1; HAP1  
 A:Cross-references: GDB:129088; OMIM:107748  
 A:Map position: 14q11.2-14q12  
 A:Introns: 20/1; 82/3; 147/1  
 A:Note: The list of introns may be incomplete  
 C:Function:  
 A:Description: catalyzes endonucleolytic cleavage near apurinic or apyrimidinic sites to  
 A:Pathway: DNA repair  
 C:Superfamily: exodeoxyribonuclease III  
 C:Keywords: carbon-oxygen lyase; DNA repair; endonuclease; magnesium; nucleus  
 F:2-318/Product: DNA (apurinic or apyrimidinic site) lyase I #status experimental <MAT>  
 F:96/Binding site: magnesium (Glu) #status predicted  
 F:309/Active site: His #status predicted

Query Match 100.0%; Score 78; DB 2; Length 318;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAGEPALYEDPPD 14  
 |||||||  
 DB 37 AAGEPALYEDPPD 50

RESULT 2

S42397  
 DNA (apurinic or apyrimidinic site) lyase (EC 4.2.99.18) - rat  
 N:Alternate names: apurinic/aprimidinic endonuclease  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 20-Oct-1994 #sequence, revision 10-Nov-1995 #text, change 18-Feb-2000  
 C:Accession: S42397  
 R:Willson, T.M.; Carney, J.P.; Kelley, M.R.  
 Nucleic Acids Res. 22, 530-531, 1994  
 A:Title: Cloning of the multifunctional rat apurinic/aprimidinic endonuclease (RAPEN)  
 A:Reference number: S42397; MUID:94173709; PMID:7510394  
 A:Accession: S42397  
 A:Molecule type: mRNA  
 A:Residues: 1-316 <WIL>  
 A:Cross-references: GB:L27076; NID:9468370; PIDN:AAA21019.1; PID:9468371  
 C:Superfamily: exodeoxyribonuclease III  
 C:Keywords: carbon-oxygen lyase; DNA repair; endonuclease; nucleus

Query Match 94.9%; Score 74; DB 2; Length 316;  
 Best Local Similarity 92.9%; Pred. No. 5.8e-05;  
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AAGEPALYEDPPD 14  
 |||||||  
 DB 35 AAGEPALYEDPPD 48

RESULT 3  
 A39500  
 DNA (apurinic or apyrimidinic site) lyase (EC 4.2.99.18) - mouse  
 N:Alternate names: apurinic/aprimidinic endonuclease; deoxyribonuclease  
 C:Species: Mus musculus (house mouse)  
 C:Date: 30-Jun-1992 #sequence, revision 30-Jun-1992 #text, change 20-Jun-2000  
 C:Accession: A39500; S17524; A56747; I49098; J00372  
 R:Seki, S.; Akiyama, K.; Watanabe, S.; Hataushika, M.; Ikeda, S.; Tsutsui, K.  
 J. Biol. Chem. 266, 20797-20802, 1991  
 A:Title: cDNA and deduced amino acid sequence of a mouse DNA repair enzyme (APEX nucle  
 A:Reference number: A39500; MUID:92041936; PMID:1939131  
 A:Accession: A39500  
 A:Molecule type: mRNA  
 A:Residues: 1-317 <SEK>  
 A:Cross-references: GB:D90374; NID:9220336; PIDN:BA01382.1; PID:9220337  
 A:Note: Part of this sequence, including the amino end of the mature protein, was con  
 R:Seki, S.; Ikeda, S.; Watanabe, S.; Hataushika, M.; Tsutsui, K.; Akiyama, K.; Zhang,  
 Biochim. Biophys. Acta 1079, 57-64, 1991  
 A:Title: A mouse DNA repair enzyme (APEX nuclease) having exonuclease and apurinic/ap  
 A:Reference number: S17524; MUID:91363416; PMID:1716153  
 A:Accession: S17524  
 A:Molecule type: protein  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 2-18, 'XXX', 22 <SE2>  
 R:Akiyama, K.; Nagao, K.; Oshida, T.; Tsutsui, K.; Yoshida, M.C.; Seki, S.  
 Genomics 26, 63-69, 1995  
 A:Title: Cloning, sequence analysis, and chromosomal assignment of the mouse Apex gen  
 A:Reference number: A56747; MUID:95301294; PMID:7782087  
 A:Accession: A56747  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-18 <AKT>  
 A:Cross-references: GB:D38077  
 R:Takiguchi, Y.; Chen, D.J.  
 Mamm. Genome 5, 717-722, 1994  
 A:Title: Genomic structure of the mouse apurinic/aprimidinic endonuclease gene.  
 A:Reference number: I49098; MUID:95178846; PMID:7533013  
 A:Accession: I49098  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-317 <RES>  
 A:Cross-references: EMBL:U12273; NID:9533107; PIDN:AA013769.1; PID:9533108  
 C:Genetics:  
 A:Gene: MGI:Apex  
 A:Cross-references: MGI:88042  
 A:Introns: 19/1; 81/3; 146/1  
 C:Superfamily: exodeoxyribonuclease III

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2003, 14:36:28 ; Search time 16 Seconds

(without alignments)  
84.118 Million cell updates/sec

Title: US-10-001-426-2

Sequence: 1 AAGGAPALYEDPPD 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR-73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	78	100.0	318	2	S23550	DNA-(apurinic or a
2	74	94.9	316	2	S42397	DNA-(apurinic or a
3	74	94.9	317	2	A39500	DNA-(apurinic or a
4	62	79.5	318	2	S26830	DNA-(apurinic or a
5	44	56.4	328	2	A13600	UDPglucose 4-epime
6	43.5	55.8	612	2	C72293	hypothetical prote
7	43	55.1	3104	2	S20473	fatty-acid synthas
8	42	53.8	303	2	T42699	hypothetical prote
9	42	53.8	1153	2	A49676	nitric-oxide synth
10	42	53.8	1182	2	I48378	hairless protein -
11	42	53.8	1453	2	F88640	protein F52C12.4 l
12	41	52.6	126	2	A84567	hypothetical prote
13	41	52.6	293	2	AP0617	probable DNA methy
14	41	52.6	1144	1	A43271	nitric-oxide synth
15	41	52.6	1147	1	S47647	nitric-oxide synth
16	41	52.6	1147	1	I56575	nitric-oxide synth
17	41	52.6	1147	1	S38253	nitric-oxide synth
18	41	52.6	1147	1	I53165	nitric-oxide synth
19	41	52.6	1147	1	JC5027	nitric-oxide synth
20	41	52.6	1147	2	JC5028	nitric-oxide synth
21	41	52.6	1147	2	JC5029	nitric-oxide synth
22	41	52.6	1147	2	JC5030	nitric-oxide synth
23	40	51.3	95	2	S65440	nitric-oxide synth
24	40	51.3	201	2	H70615	hypothetical prote
25	40	51.3	219	2	A45384	GTP-binding protei
26	40	51.3	369	2	T46950	probable ATPase co
27	40	51.3	450	2	DB7342	multidrug resistan
28	40	51.3	586	1	A43675	capsid protein p40
29	40	51.3	646	2	T28868	hypothetical prote

30	40	51.3	687	2	H72485	probable hydantoin
31	40	51.3	825	2	A48537	starch branching e
32	40	51.3	924	2	E87092	probable ribonucle
33	40	51.3	953	2	B70681	probable rne prote
34	40	51.3	987	2	G98211	hypothetical prote
35	40	51.3	987	2	AB3075	sarcosine oxidase
36	40	51.3	1094	2	C59438	KIAA1688 protein l
37	40	51.3	1123	2	A72311	conserved hypotet
38	40	51.3	1354	2	T13930	tripeptidyl-peptid
39	40	51.3	138	2	S24107	envelope protein -
40	39	50.0	138	2	S24107	competence protein
41	39	50.0	241	2	C82852	hypothetical prote
42	39	50.0	267	2	B84089	hypothetical prote
43	39	50.0	268	2	F70823	prolactin receptor
44	39	50.0	608	2	I53269	heat shock protein
45	39	50.0	717	1	HHFP83	hypothetical prote
			4077	2	T17484	

## ALIGNMENTS

RESULT 1  
DNA-(apurinic or apyrimidinic site) lyase (EC 4.2.99.18) 1 precursor - human  
N:Alternate names: Ap endonuclease 1 (APEN); APL; APEX nuclease; apurinic/aprimidini  
C:Species: Homo sapiens (man)  
C:Date: 22-Nov-1993 #sequence revision 10-Nov-1995 #text change 20-Jun-2000  
C:Accession: S23550; S25570; S35456; S23935; S47521; I39472; S34422; A41631;  
R:Xanthoudakis, S.; Miao, G.; Wang, F.; Fan, Y.C.E.; Curran, T.  
EMBO J. 11, 3323-3335, 1992  
A:Title: Redox activation of Fos-Jun DNA binding activity is mediated by a DNA repair  
A:Reference number: S23550; MUID:92371440; PMID:1380454  
A:Accession: S23550  
A:Molecule type: mRNA; protein  
A:Residues: 1-318 <XAN>  
A:Cross-references: GB:S43127; NID:9254068; PIDN:AAB22977.1; PID:9254069  
A:Note: part of this sequence, including the amino end of the mature protein, was det  
R:Robson, C.N.; Hochhauser, D.; Craig, R.; Rack, K.; Buckle, V.D.; Hickson, I.D.  
Nucleic Acids Res. 20, 4417-4421, 1992  
A:Title: Structure of the human DNA repair gene HAP1 and its localisation to chromoso  
A:Reference number: S25570; MUID:93027134; PMID:1383925  
A:Accession: S25570  
A:Molecule type: DNA  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Residues: 1-318 <ROB>  
A:Cross-references: EMBL:X66133; NID:932021; PIDN:CNA46925.1; PID:932022  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1992  
R:Cheng, X.B.; Bunville, J.; Patterson, T.A.  
Nucleic Acids Res. 20, 370, 1992  
A:Title: Nucleotide sequence of a cDNA for an apurinic/aprimidinic endonuclease from  
A:Reference number: S35454; MUID:92158631; PMID:1371347  
A:Accession: S35454  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-318 <CHN>  
A:Cross-references: EMBL:M81955; NID:9178746; PIDN:AAA58372.1; PID:9178747  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1992  
R:Zhao, B.; Grandy, D.K.; Hagerup, J.M.; Magenis, R.E.; Smith, L.; Chauman, B.C.; Hen  
Nucleic Acids Res. 20, 4097-4098, 1992  
A:Title: The human gene for apurinic/aprimidinic endonuclease (HAP1): sequence and  
A:Reference number: S35456; MUID:92375705; PMID:1380694  
A:Accession: S35456  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-318 <HNA>  
A:Cross-references: EMBL:M92444; NID:9183779; PIDN:AAA56629.1; PID:9183780  
R:Seiki, S.; Hattushika, M.; Watanabe, S.; Akiyama, K.; Nagao, K.; Tsutsui, K.  
Biochim. Biophys. Acta 1131, 287-299, 1992  
A:Title: CDNA cloning, sequencing, expression and possible domain structure of human  
A:Reference number: S23935; MUID:92329542; PMID:1627644  
A:Accession: S23935  
A:Status: preliminary  
A:Molecule type: mRNA

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2003, 14:23:53 : Search time 11 Seconds

(without alignments)  
52.788 Million cell updates/sec

Title: US-10-001-426-2

Sequence: 1 AAGEPALVEDPPD 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	100.0	317	1	PEP1_HUMAN
2	74	94.9	316	1	PEP1_MOUSE
3	74	94.9	316	1	PEP1_RAT
4	62	79.5	317	1	PEP1_BOVIN
5	42	53.8	215	1	NOS2_PTC
6	42	53.8	679	1	TKT1_KTULA
7	42	53.8	746	1	STM2_HUMAN
8	42	53.8	1149	1	NS2_CAVPO
9	42	53.8	1153	1	NS2A_HUMAN
10	42	53.8	1181	1	HAIR_RAT
11	42	53.8	1182	1	HAIR_MOUSE
12	41	52.6	130	1	EXR3_ARATH
13	41	52.6	1144	1	NOS2_MOUSE
14	41	52.6	1147	1	NOS2_RAT
15	41	52.6	1147	1	NS2D_HUMAN
16	40	51.3	209	1	CXX1_HUMAN
17	40	51.3	219	1	RB3D_MOUSE
18	40	51.3	219	1	RB3D_RAT
19	40	51.3	566	1	VP40_ILVT
20	40	51.3	591	1	NOS2_CANFA
21	40	51.3	733	1	HIC1_HUMAN
22	40	51.3	892	1	HIC1_MOUSE
23	40	51.3	987	1	SOXA_RHME
24	39.5	50.6	561	1	KCN1_HUMAN
25	39	50.0	268	1	Y739_MYCTU
26	39	50.0	351	1	KLF2_RAT
27	39	50.0	354	1	KLF2_MOUSE
28	39	50.0	608	1	PRLR_MOUSE
29	39	50.0	716	1	HSB3_DROAV
30	39	50.0	717	1	HSB3_DROME
31	39	50.0	841	1	AXN_CHICK
32	39	50.0	857	1	DD24_MOUSE
33	38	48.7	390	1	UGAT_MOUSE

## ALIGNMENTS

RESULT 1	APPEL_HUMAN	STANDARD:	PRT:	317 AA.
AC	P27695: Q99775: Q969L5:			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	DNA-(apurinic or apyrimidinic site) lyase (EC 4.2.99.16) (AP endonuclease 1) (APEX nuclease) (APEN) (REF-1 protein).			
GN	APEX OR APE OR HAP1 OR REPI OR APX.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID:9606:			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-93027134; PubMed-1383925;			
RA	Robson C.N., Hochhauser D., Craig R., Rack K., Buckle I.D.,			
RA	Hickson I.D.;			
RT	"Structure of the human DNA repair gene HAP1 and its localisation to chromosome 14q 11.2-12.";			
RL	Nucleic Acids Res. 20:4417-4421(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-Melanocyte;			
RX	MEDLINE-92051291; PubMed-1719477;			
RA	Robson C.N., Hickson I.D.;			
RT	"Isolation of cDNA clones encoding a human apurinic/apyrimidinic endonuclease that corrects DNA repair and mutagenesis defects in E. coli xth (exonuclease III) mutants.";			
RL	Nucleic Acids Res. 19:5519-5523(1991).			
RN	[3]			
RP	SEQUENCE FROM N.A. AND SEQUENCE OF 1-20.			
RX	MEDLINE-92371440; PubMed-1380454;			
RA	Xanthoudakis S., Mao G., Wang F., Pan Y.-C.E., Curran T.;			
RT	"Redox activation of Fos-Jun DNA binding activity is mediated by a DNA repair enzyme.";			
RL	EMBO J. 11:3323-3335(1992).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-92375705; PubMed-1380694;			
RA	Zhao B., Grandy D.K., Hagerup J.M., Magenis R.E., Smith L.,			
RA	Chauban B.C., Henner W.D.;			
RT	"The human gene for apurinic/apyrimidinic endonuclease (HAP1): sequence and localization to chromosome 14 band q12.";			
RL	Nucleic Acids Res. 20:4097-4098(1992).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-92329542; PubMed-1627644;			
RA	Seki S., Hatsuoka M., Watanabe S., Akiyama K., Tsutsui K.;			
RT	"cDNA cloning, sequencing, expression and possible domain structure of human APEX nuclease homologous to Escherichia coli exonuclease III.";			
RL	Biochim. Biophys. Acta 1131:287-299(1992).			
RN	[6]			
RP	SEQUENCE FROM N.A.			

34	38	48.7	446	1	HOSM_YARLI
35	38	48.7	456	1	GLMT_ECOLI
36	38	48.7	456	1	GLMT_HAEIN
37	38	48.7	576	1	ACH2_DROME
38	38	48.7	933	1	PRGR_HUMAN
39	38	48.7	1111	1	GLI1_MOUSE
40	38	48.7	1136	1	NOS2_CHICK
41	38	48.7	1322	1	PUR4_XYLF
42	38	48.7	1447	1	DCC_HUMAN
43	38	48.7	1447	1	DCC_MOUSE
44	37	47.4	198	1	RB16_RAT
45	37	47.4	296	1	Y121_HUMAN

012726	yarrowia 11
P17114	escherichia
P43889	haemophilus
P17644	drosophila
P06401	homo sapien
P47806	mus musculu
Q90703	gallus galli
Q99d66	xyloella fas
P43146	homo sapien
P70211	mus musculu
P35291	rattus norv
Q14135	homo sapien

RX MEDLINE=94368844; PubMed=8086453;  
 RA Akiyama K., Seki S., Oshida T., Yoshida M.;  
 RT "Structure, promoter analysis and chromosomal assignment of the human  
 RT Apex gene";  
 RL Biochim. Biophys. Acta 1219:15-25(1994).  
 RN [17]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=92107968; PubMed=1722334;  
 RT Demple B., Herman T., Chen D.S.;  
 RT "Cloning and expression of APE, the cDNA encoding the major human  
 RT apurinic endonuclease: definition of a family of DNA repair  
 RT enzymes";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:11450-11454(1991).  
 RN [18]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92158631; PubMed=1371347;  
 RA Cheng X.B., Bunville J., Patterson T.A.;  
 RT "Nucleotide sequence of a cDNA for an apurinic/aprimidinic  
 RT endonuclease from HeLa cells";  
 RL Nucleic Acids Res. 20:370-370(1992).  
 RN [19]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=97264341; PubMed=9110174;  
 RA Yu W., Andersson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,  
 RA Ricardente J.T., Wentland K.A., Lennon G., Gibbs R.A.;  
 RT "Large-scale concatenation cDNA sequencing";  
 RL Genome Res. 7:353-358(1997).  
 RN [110]  
 RP SEQUENCE FROM N.A., AND VARIANT GIU-147.  
 RC TISSUE=Brain, Lung, and Skin;  
 RA Strassberg R.;  
 RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 RL [111]  
 RP SEQUENCE OF 1-145 FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=93258307; PubMed=1284593;  
 RA Harrison L., Ascione G., Menninger J.C., Ward D.C., Demple B.;  
 RT "Human apurinic endonuclease gene (APE): structure and genomic  
 RT mapping (chromosome 14q11.2-12).";  
 RL Hum. Mol. Genet. 1:677-680(1992).  
 RN [112]  
 RP MUTAGENESIS OF ASN-211.  
 RX MEDLINE=97086686; PubMed=8932375;  
 RA Rothwell D.G., Hickson I.D.;  
 RT "Asparagine 212 is essential for abasic site recognition by the human  
 RT DNA repair endonuclease HAP1";  
 RL Nucleic Acids Res. 24:4217-4221(1996).  
 RN [113]  
 RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).  
 RX MEDLINE=21184546; PubMed=11286553;  
 RA Beerlink P.T., Segelke B.W., Hadi M.Z., Erzberger J.P.,  
 RA Wilson D.M. III, Rupp B.;  
 RT "Two divalent metal ions in the active site of a new crystal form of  
 RT human apurinic/aprimidinic endonuclease, apei: implications for the  
 RT catalytic mechanism";  
 RL J. Mol. Biol. 307:1023-1034(2001).  
 RN [114]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 31-317.  
 RX MEDLINE=98063001; PubMed=9351835;  
 RA Gorman M.A., Morera S., Rothwell D.G., de la Fortelle E., Mol C.D.,  
 RA Tainer J.A., Hickson I.D., Freemont P.S.;  
 RT "The crystal structure of the human DNA repair endonuclease HAP1  
 RT suggests the recognition of extra-helical deoxyribose at DNA abasic  
 RT sites";  
 RL EMBO J. 16:6548-6558(1997).  
 RN [115]  
 RP X-RAY CRYSTALLOGRAPHY (2.65 ANGSTROMS) OF 39-317.  
 RX MEDLINE=20129262; PubMed=10667800;  
 RA Mol C.D., Izumi T., Mitra S., Tainer J.A.;  
 RT "DNA-bound structures and mutants reveal abasic DNA binding by APE1  
 RT DNA repair and coordination.";

RL Nature 403:451-456(2000).  
 CC -1- FUNCTION: REPAIRS OXIDATIVE DNA DAMAGES IN VITRO. MAY HAVE A ROLE  
 CC IN PROTECTION AGAINST CELL LETHALITY AND SUPPRESSION OF MUTATIONS.  
 CC REMOVES THE BLOCKING GROUPS FROM THE 3' TERMINI OF THE DNA STRAND  
 CC BREAKS GENERATED BY IONIZING RADIATIONS AND BLEOMYCIN.  
 CC -1- CATALYTIC ACTIVITY: The C-O-P bond 3' to the apurinic or  
 CC apyrimidinic site in DNA is broken by a beta-elimination reaction,  
 CC leaving a 3'-terminal unsaturated sugar and a product with a  
 CC terminal 5'-phosphate.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SIMILARITY: BELONGS TO THE AP/EXO FAMILY OF DNA REPAIR ENZYMES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X66133; CAA46925.1; -  
 CC EMBL: X59764; CAA42437.1; -  
 CC EMBL: M92444; AAA58629.1; -  
 CC EMBL: M99703; AAA58373.1; -  
 CC EMBL: D90373; BAA14381.1; -  
 CC EMBL: D13370; BAA02633.1; -  
 CC EMBL: M80261; AAA58371.1; -  
 CC EMBL: M81955; AAA58372.1; -  
 CC EMBL: U79268; AAB50212.1; -  
 CC EMBL: BC002338; AAH02338.1; -  
 CC EMBL: BC004979; AAH04979.1; -  
 CC EMBL: BC008145; AAH08145.1; -  
 CC EMBL: BC019291; AAH19291.1; -  
 CC EMBL: S43127; AAB22977.1; -  
 CC PIR: A41631; A41631.  
 CC PIR: JU0373; JU0373.  
 CC PIR: S22072; S22072.  
 CC PIR: S23550; S23550.  
 CC PIR: S23935; S23935.  
 CC PIR: S25570; S25570.  
 CC PIR: S35456; S35456.  
 CC PIR: S35454; S35454.  
 CC PDB: 1BX; 22-JUN-99.  
 CC PDB: 1E9N; 16-FEB-01.  
 CC PDB: 1HD7; 16-FEB-01.  
 CC PDB: 1DE8; 02-FEB-00.  
 CC PDB: 1DE9; 02-FEB-00.  
 CC PDB: 1DEW; 02-FEB-00.  
 CC TRANSFAC: T04878; -  
 CC GeneW: HGNC:587; APEX.  
 CC MIM: 107748; -  
 CC InterPro: IPR000097; ApEndonuclease.  
 CC InterPro: IPR004442; ExoDNase\_III.  
 CC InterPro: IPR004808; ExoIII\_xth.  
 CC InterPro: IPR005135; Exo\_endo\_phos.  
 CC Pfam: PF03372; Exo\_endo\_phos; 1.  
 CC TIGRfam: TIGR00195; exoDNase\_III; 1.  
 CC TIGRfam: TIGR00633; xth; 1.  
 CC PROSITE: PS00726; AP\_NUCLEASE\_F1\_1; 1.  
 CC PROSITE: PS00727; AP\_NUCLEASE\_F1\_2; 1.  
 CC PROSITE: PS00728; AP\_NUCLEASE\_F1\_3; 1.  
 CC DNA repair; Lyase; Nuclear protein; Polymorphism; 3D-structure.  
 CC INIT\_MET 0 0  
 CC METAL 67 67 MAGNESIUM OR MANGANESE.  
 CC METAL 95 95 MAGNESIUM OR MANGANESE.  
 CC METAL 209 209 MAGNESIUM OR MANGANESE.  
 CC METAL 211 211 MAGNESIUM OR MANGANESE.  
 CC METAL 307 307 MAGNESIUM OR MANGANESE.  
 CC METAL 308 308 MAGNESIUM OR MANGANESE.  
 CC ACT\_SITE 308 308 GENERAL BASE (BY SIMILARITY).  
 CC VARIANT 50 50 O -> H (IN DBSNP:1048945).  
 CC /FTId=VAR\_013455.



Query Match 100.0%: Score 78: DB 1: Length 317;  
 Best Local Similarity 100.0%: Pred. No. 6.3e-06;  
 Matches 14: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY 1 AAGEGPALYEDPPD 14  
 |||||||  
 DB 36 AAGEGPALYEDPPD 49

RESULT 2  
 APEL\_MOUSE STANDARD: PRT: 316 AA.  
 AC P28352:  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE DNA-(apurinic or apyrimidinic site) lyase (EC 4.2.99.18) (AP  
 DE endonuclease 1) (APEX nuclease) (APEN).  
 GN APEX OR APE.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-92041936; PubMed-1939131;  
 RA Seki S., Akiyama K., Watanabe S., Hatsuhashiki M., Ikeda S.,  
 RA Tsutsui K.;  
 RT "cDNA and deduced amino acid sequence of a mouse DNA repair enzyme  
 RT (APEX nuclease) with significant homology to Escherichia coli  
 RT exonuclease III".  
 RL J. Biol. Chem. 266:20797-20802(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-129; TISSUE=Embryo;  
 RC MEDLINE-95178846; PubMed-7533013;  
 RA Takiguchi Y., Chen D.J.;  
 RT "Genomic structure of the mouse apurinic/apyrimidinic endonuclease  
 RT gene".  
 RL Mamm. Genome 5:717-722(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-BALB/C; TISSUE=Blood;  
 RC MEDLINE-95301294; PubMed-7782087;  
 RA Akiyama K., Nagao K., Oshida T., Tsutsui K., Yoshida M.C., Seki S.;  
 RT "Cloning, sequence analysis, and chromosomal assignment of the mouse  
 RT ApeX gene".  
 RL Genomics 26:63-69(1995).  
 RN [4]  
 RP PARTIAL SEQUENCE OF 1-21. AND CHARACTERIZATION.  
 RC TISSUE=Ascltes;  
 RA MEDLINE-91163416; PubMed-1716153;  
 RA Seki S., Ikeda S., Watanabe S., Hatsuhashiki M., Tsutsui K., Akiyama K.,  
 RA Zhang B.;  
 RT "A mouse DNA repair enzyme (APEX nuclease) having exonuclease and  
 RT apurinic/apyrimidinic endonuclease activities: purification and  
 RT characterization".  
 RL Biochim. Biophys. Acta 1079:57-64(1991).  
 CC -1- FUNCTION: REPAIRS OXIDATIVE DNA DAMAGES IN VITRO. MAY HAVE A ROLE  
 CC IN PROTECTION AGAINST CELL LETHALITY AND SUPPRESSION OF MUTATIONS.  
 CC REMOVES THE BLOCKING GROUPS FROM THE 3' TERMINI OF THE DNA STRAND  
 CC BREAKS GENERATED BY IONIZING RADIATIONS AND BLEOMYCIN.  
 CC -1- CATALYTIC ACTIVITY: The C-O-P bond 3' to the apurinic or  
 CC apyrimidinic site in DNA is broken by a beta-elimination reaction,  
 CC leaving a 3'-terminal unsaturated sugar and a product with a  
 CC terminal 5'-phosphate.  
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: BELONGS TO THE AP/EXOA FAMILY OF DNA REPAIR ENZYMES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: D90374; BAA14382.1; -  
 DR EMBL: U12273; AAC13769.1; -  
 DR EMBL: D38077; BAA07270.1; -  
 DR PIR: S17524; S17524.  
 DR PIR: A39500; A39500.  
 DR HSSP: P27695; 1BX.  
 DR MGD: MGI:88042; Apex.  
 DR InterPro: IPR000097; APendonc1se1.  
 DR InterPro: IPR004442; EXODNase\_III.  
 DR InterPro: IPR004808; EXOIII\_xth.  
 DR InterPro: IPR005135; Exo\_endo\_phos.  
 DR Pfam: PF03172; Exo\_endo\_phos. 1.  
 DR TIGRFAMS: TIGR00195; exodnase\_III. 1.  
 DR TIGRFAMS: TIGR00633; xth. 1.  
 DR PROSITE: PS00726; AP\_NUCLEASE\_F1\_1; 1.  
 DR PROSITE: PS00727; AP\_NUCLEASE\_F1\_2; 1.  
 DR PROSITE: PS00728; AP\_NUCLEASE\_F1\_3; 1.  
 KW DNA repair; Lyase; Nuclear protein.  
 FT INIT\_MER 0  
 FT METAL 66 66 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 FT METAL 94 94 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 FT METAL 208 208 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 FT METAL 210 210 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 FT METAL 306 306 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 FT METAL 307 307 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 FT SITE 210 210 IMPORTANT FOR SUBSTRATE RECOGNITION (BY  
 FT SIMILARITY).  
 FT ACT\_SITE 307 307 GENERAL BASE (BY SIMILARITY).  
 FT SQ SEQUENCE 316 AA; 35359 MW; CF8999FA8182AA558 CRC64;

Query Match 94.9%: Score 74: DB 1: Length 316;  
 Best Local Similarity 92.9%: Pred. No. 3e-05;  
 Matches 13: Conservative 0: Mismatches 1: Indels 0: Gaps 0;

OY 1 AAGEGPALYEDPPD 14  
 |||||||  
 DB 35 AAGEGPALYEDPPD 48

RESULT 3  
 APEL\_RAT STANDARD: PRT: 316 AA.  
 AC P43138:  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE DNA-(apurinic or apyrimidinic site) lyase (EC 4.2.99.18) (AP  
 DE endonuclease 1) (APEX nuclease) (APEN).  
 GN APEX OR APE.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-Sprague-Dawley; TISSUE=Testis;  
 RC MEDLINE-94173709; PubMed-7510394;  
 RA Wilson T.M., Carney J.P., Kelley M.R.;  
 RT "Cloning of the multifunctional rat apurinic/apyrimidinic  
 RT endonuclease (APEN)/Redox factor from an immature T cell line".  
 RL Nucleic Acids Res. 22:530-531(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-Sprague-Dawley; TISSUE=Brain;  
 RC Tan Y., Akiyama K., Seki S., Tabayashi T., Tanjima M.;  
 RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.  
 RN [3]

RP SEQUENCE FROM N.A.  
 RA Yao M., Akiyama K., Tan Y., Sarker A.H., Ikeda S., Alam S.S.,  
 RA Tetsui K., Yoshida M.C., Seki S.,  
 RT "Genomic structure of the rat Apex (major AP endonuclease) gene with  
 RT an adjacent putative O-sialoglycoproteinase (Prsmj/Gcp11) gene and a  
 RT processed pseudogene (Apep1).";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: REPAIRS OXIDATIVE DNA DAMAGES IN VITRO. MAY HAVE A ROLE  
 CC IN PROTECTION AGAINST CELL LETHALITY AND SUPPRESSION OF MUTATIONS.  
 CC REMOVES THE BLOCKING GROUPS FROM THE 3' TERMINI OF THE DNA STRAND  
 CC BREAKS GENERATED BY IONIZING RADIATIONS AND BLEOMYCIN.  
 CC -1- CATALYTIC ACTIVITY: The C-O-P bond 3' to the apurinic or  
 CC apyrimidinic site in DNA is broken by a beta-elimination reaction,  
 CC leaving a 3'-terminal unsaturated sugar and a product with a  
 CC terminal 5'-phosphate.  
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: BELONGS TO THE AP/EXO FAMILY OF DNA REPAIR ENZYMES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L27076; AAA21019.1; -;  
 CC EMBL: D44495; BAA07938.1; -;  
 CC EMBL: AB023065; BAA82124.1; -;  
 CC HSSP: P27695; 1BIY.  
 CC InterPro: IPR000097; Appendoc1sel.  
 CC InterPro: IPR004442; ExoDNase\_III.  
 CC InterPro: IPR004808; ExoIII\_xth.  
 CC InterPro: IPR005135; Exo\_endo\_phos.  
 CC Pfam: PF03372; Exo\_endo\_phos; 1.  
 CC TIGRfams: TIGR00195; exoDNase\_III; 1.  
 CC TIGRfams: TIGR00633; xth; 1.  
 CC PROSITE: PS00726; AP\_NUCLEASE\_FL\_1; 1.  
 CC PROSITE: PS00727; AP\_NUCLEASE\_FL\_2; 1.  
 CC PROSITE: PS00728; AP\_NUCLEASE\_FL\_3; 1.  
 CC KW DNA repair; Lyase; Nuclear protein.  
 CC FT INIT MET 0  
 CC FT METAL 66 66 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 CC FT METAL 94 94 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 CC FT METAL 208 208 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 CC FT METAL 210 210 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 CC FT METAL 306 306 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 CC FT METAL 307 307 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 CC FT ACT SITE 307 307 GENERAL BASE (BY SIMILARITY).  
 CC FT CONFLICT 9 9 MISSING (IN REF. 1).  
 CC FT CONFLICT 235 235 R -> A (IN REF. 1).  
 CC FT CONFLICT 287 287 H -> Q (IN REF. 1).  
 CC FT CONFLICT 287 287  
 CC SEQUENCE 316 AA; 35407 MW; FB27D005917C4116 CRC64;  
 Query Match 94.9%; Score 74; DB 1; Length 316;  
 Best Local Similarity 92.9%; Pred. No. 3e-05;  
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAGEGALYEDPPD 14  
 DB 35 AAGEGALYEDPPD 48  
 RESULT 4  
 ID APEL\_BOVIN STANDARD; PRT; 317 AA.  
 AC P23196;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DNA-(apurinic or apyrimidinic site) lyase (EC 4.2.99.18) (AP  
 endonuclease 1) (ApeX nuclease) (APEN).

GN APEX OR APE OR BAP1.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovine; Bos.  
 CC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-18.  
 RC TISSUE=Thymus;  
 RX MEDLINE=91212169; Pubmed=1708495;  
 RA ROSSON C.N., Milne A.M., Pappin D.J.C., Hickson I.D.;  
 RT "Isolation of cDNA clones encoding an enzyme from bovine cells that  
 RT repairs oxidative DNA damage in vitro: homology with bacterial repair  
 RT enzymes";  
 RL Nucleic Acids Res. 19:1087-1092(1991).  
 RN [2]  
 RP SEQUENCE OF 1-22.  
 RC TISSUE=Thymus;  
 RX MEDLINE=87289027; Pubmed=2441359;  
 RA Henner W.D., Klier N.P., Jorgensen T.J., Munck J.-N.;  
 RT "Purification and amino-terminal amino acid sequence of an  
 RT apurinic/apyrimidinic endonuclease from calf thymus";  
 RL Nucleic Acids Res. 15:5529-5544(1987).  
 CC -1- FUNCTION: REPAIRS OXIDATIVE DNA DAMAGES IN VITRO. MAY HAVE A ROLE  
 CC IN PROTECTION AGAINST CELL LETHALITY AND SUPPRESSION OF MUTATIONS.  
 CC REMOVES THE BLOCKING GROUPS FROM THE 3' TERMINI OF THE DNA STRAND  
 CC BREAKS GENERATED BY IONIZING RADIATIONS AND BLEOMYCIN.  
 CC -1- CATALYTIC ACTIVITY: The C-O-P bond 3' to the apurinic or  
 CC apyrimidinic site in DNA is broken by a beta-elimination reaction,  
 CC leaving a 3'-terminal unsaturated sugar and a product with a  
 CC terminal 5'-phosphate.  
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: THYMUS.  
 CC -1- INDUCTION: BY SEVERAL DNA DAMAGING AGENTS.  
 CC -1- SIMILARITY: BELONGS TO THE AP/EXO FAMILY OF DNA REPAIR ENZYMES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X56685; CAA40014.1; -;  
 CC PIR: S26830; S26830.  
 CC HSSP: P27695; 1HD7.  
 CC InterPro: IPR000097; Appendoc1sel.  
 CC InterPro: IPR004442; ExoDNase\_III.  
 CC InterPro: IPR004808; ExoIII\_xth.  
 CC InterPro: IPR005135; Exo\_endo\_phos.  
 CC Pfam: PF03372; Exo\_endo\_phos; 1.  
 CC TIGRfams: TIGR00195; exoDNase\_III; 1.  
 CC TIGRfams: TIGR00633; xth; 1.  
 CC PROSITE: PS00726; AP\_NUCLEASE\_FL\_1; 1.  
 CC PROSITE: PS00727; AP\_NUCLEASE\_FL\_2; 1.  
 CC PROSITE: PS00728; AP\_NUCLEASE\_FL\_3; 1.  
 CC KW DNA repair; Lyase; Nuclear protein.  
 CC FT INIT MET 0  
 CC FT METAL 67 67 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 CC FT METAL 95 95 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 CC FT METAL 209 209 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 CC FT METAL 211 211 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 CC FT METAL 307 307 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 CC FT METAL 308 308 MAGNESIUM OR MANGANESE (BY SIMILARITY).  
 CC FT ACT SITE 308 308 GENERAL BASE (BY SIMILARITY).  
 CC FT CONFLICT 20 21 PE -> LP (IN REF. 2).  
 CC SEQUENCE 317 AA; 35438 MW; 58A7B2029B0891B4 CRC64;  
 Query Match 79.5%; Score 62; DB 1; Length 317;  
 Best Local Similarity 78.6%; Pred. No. 0.0033;  
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AAGEGPALYEDPPD 14  
 DB 36 AAGEGPALYEDPPD 49

## RESULT 5

NOS2\_PIG

ID NOS2\_PIG STANDARD: PRT: 215 AA.

AC P79290:

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Nitric oxide synthase, inducible (EC 1.14.13.39) (NOS, type II)

DE (Inducible NOS) (1NOS) (Fragment).

GN NOS2.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

OX NCBI\_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Spleen;

RA MEDLINE=98273497; PubMed=9613441;

RA Pampusch M.S., Bennis A.M., Harsch S., Murtough M.P.;

RT Vet. Immunol. Immunopathol. 61:279-289(1998). IS A MESSENGER MOLECULE

CC WITH DIVERSE FUNCTIONS THROUGHOUT THE BODY.

CC -1- CATALYTIC ACTIVITY: L-arginine + N NADPH + M O(2) = citrulline +

CC nitric oxide + N NADP(+).

CC -1- COFACTOR: HEME. BINDS ONE MOLE EACH OF FAD AND FMN. ALSO REQUIRES

CC TETRAHYDROBIOTERIN (BH4) WHICH MAY STABILIZE THE DIMERIC FORM OF

CC THE ENZYME (BY SIMILARITY).

CC -1- ENZYME REGULATION: NOT STIMULATED BY CALCIUM/CALMODULIN (BY

CC SIMILARITY).

CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC -1- TISSUE SPECIFICITY: DETECTED IN BOTH STIMULATED AND UNSTIMULATED

CC IMMUNE CELLS AND MACROPHAGES WITH LITTLE OR NO UPREGULATION

CC FOLLOWING CELLULAR STIMULATION WITH LIPOPOLYSACCHARIDES (LPS) OR

CC CONCANAVALIN A (CONA).

CC -1- INDUCTION: LITTLE BY LPS OR CONA IN SPLEEN CELLS.

CC -1- SIMILARITY: BELONGS TO THE NOS FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

CC EMBL: U59390; AAB40614.1; -.

DR HSSP: P35228; 4NOS.

DR InterPro: IPR004030; NO\_synthase.

DR Pfam: PF02898; NO\_synthase; 1.

DR PROSITE: PS60001; NOS: PARTIAL.

KW Oxidoreductase; NADP; FAD; FMN; Calmodulin-binding; Heme;

KW Multigene family.

FT NON\_TER

FT 1

SQ SEQUENCE

215 AA; 24716 MW; FECEFSBD85B5808 CRC64;

OY 1 AAGEGPALYEDPPD 14

DB 99 AAGEGPALYEDPPD 112

## RESULT 6

TKT1\_KLUUA STANDARD: PRT: 679 AA.

AC Q12630:

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Transketolase (EC 2.2.1.1) (TK).

GN TKT1.

OS Kluyveromyces fragilis (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.

OX NCBI\_TaxID=28985;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CBS 2359/152;

RA MEDLINE=97153463; PubMed=9000376;

RA Jacoby J.J., Hehlisch J.J.;

RT "Analysis of a transketolase gene from Kluyveromyces fragilis reveals

CC that the yeast enzymes are more related to transketolases of

CC prokaryotic origins than to those of higher eukaryotes."

CC Curr. Genet. 31:15-21(1997).

CC -1- CATALYTIC ACTIVITY: Sedoheptulose 7-phosphate + D-glyceraldehyde

CC 3-phosphate -&gt; D-ribose 5-phosphate + D-xylulose 5-phosphate.

CC -1- COFACTOR: THIAMINE PYROPHOSPHATE.

CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

CC EMBL: U65983; AAB05935.1; -.

DR HSSP: P23254; TTKR.

DR InterPro: IPR000360; Transketolase.

DR Pfam: PF00456; transketolase; 1.

DR Pfam: PF02779; transketolase; 1.

DR Pfam: PF02780; transketolase; 1.

DR TIGRfam: TIGR00232; tkase\_dact; 1.

DR PROSITE: PS00801; TRANSKETOLASE\_1; 1.

DR PROSITE: PS00802; TRANSKETOLASE\_2; 1.

KW Transferrase; Thiamine pyrophosphate.

SQ SEQUENCE 679 AA; 73703 MW; BCSF3C61A78CD4A CRC64;

OY 1 AAGEGPALYEDPPD 10

DB 637 AAGEGPALYEDPPD 646

## RESULT 7

STR2\_HUMAN

ID STR2\_HUMAN

AC Q9P246; Q9P246; Q9H8R1;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Stromal interaction molecule 2 precursor.

GN STR2 OR KIAA1482.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., SEQUENCE OF 15-24, SUBUNIT, GLYCOSYLATION,

RP PHOSPHORYLATION, AND TISSUE SPECIFICITY.

RC TISSUE=Fetal brain, and Kidney;

RX MEDLINE=21356314; PubMed=11463338;  
 RA Williams R.T., Manji S.S.M., Parker N.J., Hancock M.S.,  
 RA Van Stekelenburg L., Eid J.-P., Senior P.V., Kazenwadel J.S.,  
 RA Shandela T., Saint R., Smith P.J., Dziadek M.A.;  
 "Identification and characterization of the SRIM (stromal interaction  
 RT molecule) gene family: coding for a novel class of transmembrane  
 RT proteins";  
 RL Biochem. J. 357:673-685(2001).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE-Brain;  
 RX MEDLINE=20277482; PubMed=10819331;  
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;  
 "Prediction of the coding sequences of unidentified human  
 RT genes. XVII. The complete sequences of 100 new cDNA clones from brain  
 RT which code for large proteins in vitro.";  
 RL DNA Res. 7:143-150(2000).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE-Amygdala;  
 RX MEDLINE=21154917; PubMed=11230166;  
 RA Wiemann S., Weill B., Wellenreuther R., Gassenhuber J., Glassl S.,  
 RA Ansgore W., Boecher M., Bloecher H., Bauersachs S., Blum H.,  
 RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,  
 RA Mewes H.-W., Oltensweider B., Obermaier B., Tampe J., Heubner D.,  
 RA Wambutt R., Korn B., Klein M., Poustka A.;  
 "Towards a catalog of human genes and proteins: sequencing and  
 RT analysis of 500 novel complete protein coding human cDNAs.";  
 RL Genome Res. 11:422-435(2001).  
 [4]  
 RN SEQUENCE OF 293-695 FROM N.A.  
 RP TISSUE-Ovarian carcinoma;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Magatsuna M., Hosoki T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto Y., Wakamatsu A., Nakamura Y., Nagahara K., Masuo Y.,  
 RA Niimiya K., Iwayanagi T.;  
 "NEO human cDNA sequencing project.";  
 RT Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: Possible adhesion molecule with a role in early  
 hematopoiesis by mediating attachment to stromal cells. Influences  
 CC the survival and/or proliferation of B cell precursors. Binding to  
 CC cells requires Mn(II) (By similarity).  
 CC -1- SUBUNIT: Oligomer with STIM1.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface  
 CC (Potential).  
 CC -1- TISSUE SPECIFICITY: Expressed in all tissues and tumor cell lines  
 CC examined.  
 CC -1- PTM: Glycosylated.  
 CC -1- PTM: Phosphorylated predominantly on Ser residues.  
 CC -1- SIMILARITY: CONTAINS 1 SAM DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AF328905; AAK82337.1; -;  
 DR EMBL: AB040915; BAA96006.1; ALT\_INT.  
 DR EMBL: AL136577; CAB66512.1; ALT\_INT.  
 DR EMBL: AK023369; BAB14545.1; ALT\_SEO.  
 DR InterPro: IPR001650; SAM.  
 DR SMART: SM00454; SAM; 1.  
 DR PROSITE: PS50105; SAM\_DOMAIN; 1.  
 KM Cell adhesion: Transmembrane; Coiled coil; Signal; Glycoprotein;  
 KW Phosphorylation.  
 FT SIGNAL 1 14  
 FT CHAIN 15 746 STROMAL INTERACTION MOLECULE 2.

FT DOMAIN 15 217 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 218 235 POTENTIAL.  
 FT DOMAIN 236 746 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 136 204 SAM.  
 FT DOMAIN 242 344 COILED COIL (POTENTIAL).  
 FT DOMAIN 358 394 COILED COIL (POTENTIAL).  
 FT DOMAIN 533 559 HIS/PRO-RICH.  
 FT DOMAIN 730 746 LYS-RICH.  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CONFLICT 9 9 A -> P (IN REF. 3).  
 FT CONFLICT 588 588 W -> WNTRECAAGDSQGFHVHGLVFDPDFGSYERKH  
 FT SEQUENCE 746 AA; 83971 MW; B8385557F126254 CRC64;  
 SQ  
 Query Match 53.8%; Score 42; DB 1; Length 746;  
 Best Local Similarity 66.7%; Pred. No. 21;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 OY 1 AGESEPALYEDP 12  
 DB 41 AAGDSPALMTDP 52  
 ID NOS2\_CAVPO STANDARD; PRT; 1149 AA.  
 NOS2\_CAVPO  
 RC 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Nitric oxide synthase, inducible (EC 1.14.13.39) (NOS, type II)  
 DE (Inducible NOS) (INOS).  
 GN NOS2 OR NOS.  
 OS *Cavia porcellus* (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
 OX NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hartley; TISSUE=Lung;  
 RX MEDLINE=98343971; PubMed=9677342;  
 RA Shirato M., Sakamoto T., Uchida Y., Nomura A., Ishii Y., Iijima H.,  
 RA Goto Y., Hasegawa S.;  
 RT "Molecular cloning and characterization of Ca2+-dependent inducible  
 RT nitric oxide synthase from guinea-pig lung.";  
 RL Biochem. J. 333:795-799(1998).  
 CC -1- FUNCTION: PRODUCES NITRIC OXIDE (NO) WHICH IS A MESSENGER MOLECULE  
 WITH DIVERSE FUNCTIONS THROUGHOUT THE BODY.  
 CC -1- CATALYTIC ACTIVITY: L-arginine + NADPH + M O(2) = citrulline +  
 CC nitric oxide + N NADP(+).  
 CC -1- COFACTOR: HEME. BINDS ONE MOLE EACH OF PAD AND FMN. ALSO REQUIRES  
 CC TETRAHYDROBIOTERIN (BH4) WHICH MAY STABILIZE THE DIMERIC FORM OF  
 CC THE ENZYME (BY SIMILARITY).  
 CC -1- ENZYME REGULATION: REGULATED BY CALCIUM/CALMODULIN (IN CONTRAST  
 CC WITH MOUSE NOS2).  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE LUNG AND COLON. NOT DETECTED  
 CC IN THE HEART, AORTA, LIVER, KIDNEY, AND SPLEEN.  
 CC -1- INDUCTION: BY LIPOPOLYSACCHARIDE (LPS); IN KIDNEY, SPLEEN, AND  
 CC COLON. EXPRESSION IS REDUCED IN THE PRESENCE OF LPS IN LUNG.  
 CC -1- SIMILARITY: BELONGS TO THE NOS FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AF027180; AAC31177.1; -;  
 DR EMBL: P29477; 1D07.  
 DR InterPro: IPR003097; FAD\_binding.

DR InterPro: IPR001709; FPN\_cyl\_redcse.  
 DR InterPro: IPR001094; Flavodoxin\_Like.  
 DR InterPro: IPR001226; Flavodoxin.  
 DR InterPro: IPR004030; NO\_synthase.  
 DR InterPro: IPR01433; Oxred\_FAD/NAD(P).  
 DR Pfam: PF00175; NAD\_binding\_1.  
 DR Pfam: PF00258; Flavodoxin\_1.  
 DR Pfam: PF00667; FAD\_binding\_1.  
 DR Pfam: PF02898; NO\_synthase\_1.  
 DR PRINTS: PR00369; FLAVODOXIN.  
 DR PRINTS: PR00371; FPNCR.  
 DR PROSITE: PS60001; NOS\_1.  
 DR Oxidoreductase: NADP; FAD; FMN; Calcium-binding; Calmodulin-binding;  
 KW Heme; Multigene family.  
 FT BINDING 199 199 HEME (BY SIMILARITY).  
 FT DOMAIN 308 328 CALMODULIN-BINDING (POTENTIAL).  
 FT NP\_BIND 622 653 FMN (PYRIMIDINE PART) (BY SIMILARITY).  
 FT NP\_BIND 766 777 FAD (ADP PART) (BY SIMILARITY).  
 FT NP\_BIND 902 912 FAD (FLAVIN PART) (BY SIMILARITY).  
 FT NP\_BIND 977 995 NADP (RIBOSE PART) (BY SIMILARITY).  
 FT NP\_BIND 1075 1090 NADP (ADP PART) (BY SIMILARITY).  
 SQ SEQUENCE 1149 AA; 130631 MW; 66D595A2486DB50E CRC64;

Query Match 53.8%; Score 42; DB 1; Length 1149;  
 Best Local Similarity 57.1%; Pred. No. 34;  
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 AGCGPALYEDPPD 14  
 DB 310 ADGRDELFEIPDP 323

RESULT 9  
 NS2A\_HUMAN STANDARD; PRT; 1153 AA.  
 ID NS2A\_HUMAN  
 AC P35228; Q16692; O60757; I16263; O94994;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Nitric oxide synthase, inducible (EC 1.14.13.39) (NOS, type II)  
 DE (Inducible NOS) (iNOS) (Hepatocyte NOS) (HEP-NOS).  
 GN NOS2A OR NOS2.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Liver;  
 RX MEDLINE=93234523; PubMed=7682706;  
 RA Geller D.A., Lowenstein C.J., Shapiro R.A., Nussler A.K.,  
 RA di Silvio M., Wang S.C., Nakayama D.K., Simmons R.L., Snyder S.H.,  
 RA Billiar T.R.;  
 RT "Molecular cloning and expression of inducible nitric oxide synthase  
 RT from human hepatocytes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:3491-3495(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Colorectal adenocarcinoma;  
 RX MEDLINE=9403282; PubMed=7692964;  
 RA Sherman P.A., Laubach V.E., Reep B.R., Wood E.R.;  
 RT "Purification and cDNA sequence of an inducible nitric oxide synthase  
 RT from a human tumor cell line.";  
 RL Biochemistry 32:11600-11605(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Chondrocytes;  
 RX MEDLINE=94068614; PubMed=7504305;  
 RA Charles I.G., Palmer R.M.J., Hickery M.S., Bayliss M.T.,  
 RA Chubb A.P., Hall V.S., Moss D.W., Moncada S.;  
 RT "Cloning, characterization, and expression of a cDNA encoding an  
 RT inducible nitric oxide synthase from the human chondrocyte.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:11419-11423(1993).

RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Articular chondrocytes;  
 RX MEDLINE=9436816; PubMed=7522054;  
 RA Maier R., Bilbe G., Rediske J., Lotz M.;  
 RT "Inducible nitric oxide synthase from human articular chondrocytes:  
 RT cDNA cloning and analysis of mRNA expression.";  
 RL Biochim. Biophys. Acta 1208:145-150(1994).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Retina;  
 RX MEDLINE=95091827; PubMed=7528017;  
 RA Park C.S., Pardhasaradhi K., Gianotti C., Villegas E., Krishna G.;  
 RT "Human retina expresses both constitutive and inducible isoforms of  
 RT nitric oxide synthase mRNA.";  
 RL Biochem. Biophys. Res. Commun. 205:85-91(1994).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Glioblastoma;  
 RX MEDLINE=95155267; PubMed=7531687;  
 RA Hokari A., Zeniya M., Esumi H.;  
 RT "Cloning and functional expression of human inducible nitric oxide  
 RT synthase (NOS) cDNA from a glioblastoma cell line A-172.";  
 RL J. Biochem. 116:575-581(1994).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Airway epithelium;  
 RX MEDLINE=95372368; PubMed=7544004;  
 RA Guo F.H., de Raeye R.H., Rice T.W., Stuehr D.J., Thunnissen F.B.J.M.,  
 RA Erzurum S.C.;  
 RT "Continuous nitric oxide synthetis by inducible nitric oxide synthase  
 RT in normal human airway epithelium in vivo.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:7809-7813(1995).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Cardiac myocytes;  
 RX MEDLINE=97304504; PubMed=9160867;  
 RA Luss H., Li R.-K., Shapiro R.A., Zheng E., McGowan F.X., Yonegama T.,  
 RA Nakayama K., Geller D.A., Michle D.A.G., Simmons R.L.,  
 RA Billiar T.R.;  
 RT "Dedifferentiated human ventricular cardiac myocytes express  
 RT inducible nitric oxide synthase mRNA but not protein in response to  
 RT IL-1, TNF, IFNgamma, and LPS.";  
 RL J. Mol. Cell. Cardiol. 29:1153-1165(1997).  
 RN [9]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RA Ogawa Y., Nishijima S., Goto M., Ida M.;  
 RT "Cloning and characterization of a novel splice variant of human  
 RT inducible nitric oxide synthase.";  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [10]  
 RP SEQUENCE OF 380-473 FROM N.A.  
 RC TISSUE-Kidney;  
 RX MEDLINE=95165725; PubMed=7532248;  
 RA McIay J.S., Chatterjee P., Nicolson A.G., Jardine A.G., McKay N.G.,  
 RA Ralston S.H., Grabowski P., Haltes N.E., MacLeod A.M.,  
 RA Hawksworth G.M.;  
 RT "Nitric oxide production by human proximal tubular cells: a novel  
 RT immunomodulatory mechanism?";  
 RL Kidney Int. 46:1043-1049(1994).  
 RN [11]  
 RP CHARACTERIZATION.  
 RX MEDLINE=96047340; PubMed=7558036;  
 RA Bloch K.D., Wolfgram J.R., Brown D.M., Roberts J.D., Zapol D.G.,  
 RA Lepore J.J., Filippov G., Thomas J.E., Jacob H.J., Bloch D.B.;  
 RT "Three members of the nitric oxide synthase II gene family (NOS2a,  
 RT NOS2b, and NOS2c) colocalize to human chromosome 17.";  
 RL Genomics 27:526-530(1995).  
 RN [12]  
 RP CHARACTERIZATION.  
 RX MEDLINE=98389865; PubMed=9721329;  
 RA Taylor B.S., Alarcon L.H., Billiar T.R.;  
 RT "Inducible nitric oxide synthase in the liver: regulation and

RT function.",  
 RL Biochimica 63:766-781(1998).  
 RN [13]  
 RP X-RAY CRYSTALLOGRAPHY (2.55 ANGSTROMS) OF 74-504.  
 RX MEDLINE=99340067; PubMed=10409685.  
 RA Li H., Raman C.S., Glaeser C.B., Blasko E., Young T.A., Parkinson J.F.,  
 RA Whittow M., Foulos T.L.,  
 RA "Crystal structures of zinc-free and -bound heme domain of human  
 RT inducible nitric-oxide synthase. Implications for dimer stability and  
 RT comparison with endothelial nitric-oxide synthase.";  
 RL J. Biol. Chem. 274:21276-21284(1999).  
 RP [14]  
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 82-528.  
 RX MEDLINE=99173237; PubMed=10074942;  
 RA Fischmann T.O., Hruza A., Niu X.D., Rossetta J.D., Lunn C.A.,  
 RA Dolphin E., Prongay A.J., Reichert P., Lundell D.J., Narula S.K.,  
 RA Weber P.C.;  
 RT "Structural characterization of nitric oxide synthase isoforms  
 RT reveals striking active-site conservation.";  
 RL Nat. Struct. Biol. 6:233-242(1999).  
 CC -1- FUNCTION: PRODUCES NITRIC OXIDE (NO) WHICH IS A MESSENGER MOLECULE  
 CC WITH DIVERSE FUNCTIONS THROUGHOUT THE BODY. IN MACROPHAGES, NO  
 CC MEDIATES TUMORICIDAL AND BACTERICIDAL ACTIONS.  
 CC -1- CATALYTIC ACTIVITY: L-arginine + N NADPH + M O(2) = citrulline +  
 CC nitric oxide + N NADP(+).  
 CC -1- COFACTOR: HEME. BINDS ONE MOLE EACH OF FAD AND FMN. ALSO REQUIRES  
 CC TETRAHYDROBIPTERIN (BH4) WHICH MAY STABILIZE THE DIMERIC FORM OF  
 CC THE ENZYME (BY SIMILARITY).  
 CC -1- ENZYME REGULATION: REGULATED BY CALCIUM/CALMODULIN (IN CONTRAST  
 CC WITH MOUSE NOS2). ASPIRIN INHIBITS EXPRESSION AND FUNCTION OF THIS  
 CC ENZYME AND EFFECTS MAY BE EXERTED AT THE LEVEL OF  
 CC TRANSCRIPTIONAL/POSTTRANSCRIPTIONAL MODIFICATION AND DIRECTLY ON THE  
 CC CATALYTIC ACTIVITY (BY SIMILARITY).  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are  
 CC produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE LIVER, RETINA, BONE CELLS AND  
 CC AIRWAY EPITHELIAL CELLS OF THE LUNG. NOT EXPRESSED IN THE  
 CC PLATELETS.  
 CC -1- INDUCTION: BY ENDOTOXINS AND CYTOKINES.  
 CC -1- SIMILARITY: BELONGS TO THE NOS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L09210; AAA59171.1; -;  
 DR EMBL: L24553; AAA36375.1; -;  
 DR EMBL: X73029; CAAS1512.1; -;  
 DR EMBL: U05810; AAA56666.1; -;  
 DR EMBL: U31511; AAB49041.1; -;  
 DR EMBL: D26525; BAA05531.1; -;  
 DR EMBL: U20141; AAB60366.1; -;  
 DR EMBL: AF068236; AAC19133.1; -;  
 DR EMBL: AB022318; BAA37123.1; -;  
 DR EMBL: S75615; AAD14179.1; -;  
 DR PIR: AA7475; A47475.  
 DR PDB: 1NSI; 07-JAN-00.  
 DR PDB: 2NSI; 07-JAN-00.  
 DR PDB: 4NOS; 04-FEB-00.  
 DR GeneW; HGNC:7873; NOS2A.  
 DR MIM: 163730; -;  
 DR InterPro: IPR003097; FAD\_binding.  
 DR InterPro: IPR001709; FPN\_cyl\_redctse.  
 DR InterPro: IPR001094; Flavodoxin\_like.  
 DR InterPro: IPR001226; Flavodoxin.  
 DR InterPro: IPR004030; NO\_synthase.  
 DR InterPro: IPR001433; Oxred\_FAD\_NAD(P).  
 DR Pfam: PF00175; NAD\_binding; 1.

DR Pfam: PF00258; flavodoxin; 1.  
 DR Pfam: PF00667; FAD\_binding; 1.  
 DR Pfam: PF02898; NO\_synthase; 1.  
 DR PRINTS: PR00369; FLAVODOXIN.  
 DR PRINTS: PR00371; FPNCR.  
 DR PROSITE: PS60001; NOS; 1.  
 KW Oxidoreductase; NADP; FAD; FMN; Calcium-binding; Calmodulin-binding;  
 KW Heme; Phosphorylation; Zinc; Metal-binding; Multigene family;  
 KW Alternative splicing; 3D-structure.  
 FT BINDING 200 200  
 FT DOMAIN 509 529  
 FT NP\_BIND 623 654  
 FT NP\_BIND 767 778  
 FT NP\_BIND 903 913  
 FT NP\_BIND 978 996  
 Query Match 53.8%; Score 42; DB 1; Length 1153;  
 Best Local Similarity 57.1%; Pred. No. 34;  
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 AAGEGPALEYDPPD 14  
 DB 311 ANGRDPELEFEPD 324  
 HAIR\_RAT  
 ID HAIR\_RAT STANDARD; PRT; 1181 AA.  
 AC P97609;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Hairless protein.  
 GN HR.  
 OS Rattus norvegicus (rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=Sprague-Dawley;  
 RX MEDLINE=97141510; PubMed=8987811;  
 RA Thompson C.C.;  
 RT "Thyroid hormone-responsive genes in developing cerebellum include a  
 RT novel synaptotagmin and a hairless homolog.";  
 RT J. Neurosci. 16:7832-7840(1996).  
 CC -1- FUNCTION: MAY ACT AS A TRANSCRIPTION FACTOR THAT COULD ACT ON TO  
 CC REGULATE ONE OF THE PHASES OF HAIR GROWTH.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U71293; AAC53018.1; ALT\_INIT.  
 DR InterPro: IPR003347; TF\_JmJC.  
 DR Pfam: PF02373; JmJC; 1.  
 KW Zinc-finger; DNA-binding; Nuclear protein; Transcription regulation;  
 KW Metal-binding.  
 FT ZN\_FING 594 619  
 FT SEQUENCE 1181 AA; 127307 MW; 834B7029CF8E88F0 CRC64;  
 Query Match 53.8%; Score 42; DB 1; Length 1181;  
 Best Local Similarity 63.6%; Pred. No. 35;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 GEGPALYEDPP 13  
 DB 578 GGGPATEDSP 588

```

RESULT 11
HAIR_MOUSE
ID HAIR_MOUSE STANDARD: PRT: 1182 AA.
AC 061845;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hairless protein.
GN HR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-Skin;
RX MEDLINE=94329587; PubMed=8052649;
RA Cachon-Gonzalez M.B., Fennel S., Coffin J.M., Moran C., Best S.,
RA Stoye J.P.;
RT "Structure and expression of the hairless gene of mice.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:7717-7721(1994).
CC -1- FUNCTION: MAY ACT AS A TRANSCRIPTION FACTOR THAT COULD ACT ON TO
CC REGULATE ONE OF THE PHASES OF HAIR GROWTH.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN BRAIN, HAIR
CC FOLLICLES AND INTERFOLLICULAR EPIDERMIS. NO EXPRESSION IN
CC DERMIS.
CC -1- DISEASE: HR MUTATION PRODUCES A NUMBER OF PLEIOTROPIC EFFECTS
CC INCLUDING STRUCTURAL ABNORMALITIES OF EPITHELIAL CELLS IN THE HAIR
CC FOLLICLES, HAIR LOSS TOWARDS THE END OF THE FIRST HAIR GROWTH
CC CYCLE, AND THE FAILURE OF SUBSEQUENT HAIR GROWTH CYCLES. OLDER
CC MICE CARRYING AN HR MUTATION HAVE BEEN REPORTED TO POSSESS ALTERED
CC RATIOS OF T-CELL-DEPENDENT B-CELL RESPONSES. MICE HOMozyGous FOR
CC SKIN TUMORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: Z32675; CAA83587.1; -
CC DR MGD: MGI:96223; hr.
CC DR InterPro: IPR003347; TF_JmjC.
CC DR Pfam: PF02373; JmjC; 1.
CC KM Zinc-finger; DNA-binding; Nuclear protein; Transcription regulation;
CC KM Metal-binding.
CC FT ZN_FING 620 CG-TYPE.
CC FT DOMAIN 535 540 POLY-GLY.
CC SQ SEQUENCE 1182 AA; 127182 MW; EFBBOEE62AE81B40 CRC64;
Query Match 53.8%; Score 42; DB 1; Length 1182;
Best Local Similarity 63.8%; Pred. No. 35;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 3 GEGPALYEDPP 13
ID 1:|||||111
Db 579 GCGPAMTEDSP 589

```

```

GN EXP3 OR AT2G18660 OR MSF3.4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Unsworth L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGEC).";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE EXPANSIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AC005724; M408935.1; ALT_INT.
CC DR EMBL: A7065269; AAL38745.1; -
CC DR PROSITE: PS50842; EXPANSIN_EG45; 1.
CC KM Glycoprotein; Signal; Multigene family.
CC FT SIGNAL 1 25 POTENTIAL.
CC FT CHAIN 26 130 EXPANSIN-RELATED PROTEIN 3.
CC FT DOMAIN 28 130 EXPANSIN-LIKE EG45.
CC FT CARDHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CONFLICT 130 130 MISSING (IN REF. 1).
CC SQ SEQUENCE 130 AA; 14518 MW; BOC51C605043BC9 CRC64;
Query Match 52.6%; Score 41; DB 1; Length 130;
Best Local Similarity 58.3%; Pred. No. 4; 8;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 AGECPALYEDPP 13
ID 1:|||||111
Db 25 AAGCAVYVDPP 36

```

```

RESULT 12
EXR3_ARATH
ID EXR3_ARATH STANDARD: PRT: 130 AA.
AC 092V52;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Expansin-related protein 3 precursor (Ath-ExpGamma-1.2).
DE

```

```

RESULT 13
NOS2_MOUSE
ID NOS2_MOUSE STANDARD: PRT: 1144 AA.
AC P29477; O70515; O70516;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nitric oxide synthase, inducible (EC 1.14.13.39) (NOS, type II)
DE (Inducible NOS) (iNOS) (Macrophage NOS) (MAC-NOS).
GN NOS2 OR INOSL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

RP SEQUENCE FROM N.A.  
RX MEDLINE-92229444; PubMed-1373522; Munford R.A., Swiderek K.M.,  
RA Xie Q.-W., Cho H.-J., Calaycay J.,  
RA Lee T.D., Ding A., Troso T., Nathan C.;  
RT "Cloning and characterization of inducible nitric oxide synthase from  
RT mouse macrophages.";  
RL Science 256:225-228(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-92357701; PubMed-1379716;  
RA Lowenstein C.J., Glatz C.S., Bredt D.S., Snyder S.H.;  
RT "Cloned and expressed macrophage nitric oxide synthase contrasts with  
RT the brain enzyme.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:6711-6715(1992).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-92210618; PubMed-1372907;  
RA Lyons C.R., Orloff G.J., Cunningham J.M.;  
RT "Molecular cloning and functional expression of an inducible nitric  
RT oxide synthase from a murine macrophage cell line.";  
RL J. Biol. Chem. 267:6370-6374(1992).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-96088781; PubMed-7503239;  
RA Kone B.C., Schmoebel J., Turner P., Mohaupt M.G., Cangro C.B.;  
RT "Role of NF-kappa B in the regulation of inducible nitric oxide  
RT synthase in an HTAL cell line.";  
RL Am. J. Physiol. 269:F718-F729(1995).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN-DBA/2J, BALB/CBY, SJL/J, NOD/LTJ, and B10.S/J; TISSUE=Spleen;  
RA Ma R.Z., Teuscher C.;  
RN Submitted (May-1998) to the EMBL/Genbank/DBJ databases.  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 116-496.  
RX MEDLINE-97477482; PubMed-9334294;  
RA Crane B.R., Arvai A.S., Gachhui R., Wu C., Ghosh D.K., Getzoff E.D.,  
RA Stuehr D.J., Tainer J.A.;  
RT "The structure of nitric oxide synthase oxygenase domain and  
RT inhibitor complexes.";  
RL Science 278:425-431(1997).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (2.35 ANGSTROMS) OF 77-496.  
RX MEDLINE-20031637; PubMed-10562538;  
RA Ghosh D.K., Crane B.R., Ghosh S., Wolan D., Gachhui R., Crooks C.,  
RA Presta A., Tainer J.A., Getzoff E.D., Stuehr D.J.;  
RT "Inducible nitric oxide synthase: role of the N-terminal beta-hairpin  
RT hook and pterin-binding segment in dimerization and  
RT tetrahydropterin interaction.";  
RL EMBO J. 18:6260-6270(1999).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 77-496.  
RX MEDLINE-98182450; PubMed-9516116;  
RA Crane B.R., Arvai A.S., Ghosh D.K., Wu C., Getzoff E.D., Stuehr D.J.,  
RA Tainer J.A.;  
RT "Structure of nitric oxide synthase oxygenase dimer with pterin and  
RT substrate.";  
RL Science 279:2121-2126(1998).  
RN [9]  
RP X-RAY CRYSTALLOGRAPHY (2.35 ANGSTROMS) OF 77-499.  
RX MEDLINE-20031638; PubMed-10562539;  
RA Crane B.R., Rosenfeld R.J., Arvai A.S., Ghosh D.K., Ghosh S.,  
RA Tainer J.A., Stuehr D.J., Getzoff E.D.;  
RT "N-terminal domain swapping and metal ion binding in nitric oxide  
RT synthase dimerization.";  
RL EMBO J. 18:6271-6281(1999).  
RN [10]  
RP X-RAY CRYSTALLOGRAPHY OF (2.35 ANGSTROMS) 77-496.  
RX MEDLINE-20033702; PubMed-10769116;  
RA Crane B.R., Arvai A.S., Ghosh S., Getzoff E.D., Stuehr D.J.,  
RA Tainer J.A.;  
RT "Structures of the N(omega)-hydroxy-L-arginine complex of inducible  
RT nitric oxide synthase oxygenase dimer with active and inactive

RT pterins.";  
RL Biochemistry 39:4608-4621(2000).  
RN [11]  
RP EFFECT OF ASPIRIN.  
RC TISSUE=Macrophage;  
RX MEDLINE-95372392; PubMed-7544010;  
RA Amin A.R., Vyas P., Attur M., Leszczynska-Piziak J., Patel I.R.,  
RA Weissmann G., Abramson S.B.;  
RT "The mode of action of aspirin-like drugs: effect on inducible nitric  
RT oxide synthase.";  
RL Proc. Natl. Acad. Sci. U.S.A. 92:7926-7930(1995).  
CC -!- FUNCTION: PRODUCES NITRIC OXIDE (NO) WHICH IS A MESSENGER MOLECULE  
CC WITH DIVERSE FUNCTIONS THROUGHOUT THE BODY. IN MACROPHAGES, NO  
CC MEDIATES TUMORICIDAL AND BACTERICIDAL ACTIONS.  
CC -!- CATALYTIC ACTIVITY: L-arginine + N NADPH + M O(2) -> citrulline +  
CC nitric oxide + N NADP(+).  
CC -!- COFACTOR: HEME. BINDS ONE MOLE EACH OF FAD AND FMN. ALSO REQUIRES  
CC TETRAHYDROBIPTERIN (BH4) WHICH MAY STABILIZE THE DIMERIC FORM OF  
CC THE ENZYME.  
CC -!- ENZYME REGULATION: NOT STIMULATED BY CALCIUM/CALMODULIN. ASPIRIN  
CC INHIBITS EXPRESSION AND FUNCTION OF THIS ENZYME AND EFFECTS MAY BE  
CC EXERTED AT THE LEVEL OF TRANSLATIONAL/POSTTRANSLATIONAL  
CC MODIFICATION AND DIRECTLY ON THE CATALYTIC ACTIVITY.  
CC -!- SUBUNIT: HOMODIMER.  
CC -!- TISSUE SPECIFICITY: MACROPHAGES.  
CC -!- INDUCTION: BY TREATMENT WITH ENDOTOXINS OR CYTOKINES.  
CC -!- SIMILARITY: BELONGS TO THE NOS FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
CC -----  
DR EMBL, M87039; AAA39315.1; -;  
DR EMBL, M92649; -; NOT\_ANNOTATED\_CDS.  
DR EMBL, M84373; AAA39834.1; -;  
DR EMBL, U43428; AAC52356.1; -;  
DR EMBL, AF065919; AAC17914.1; -;  
DR EMBL, AF065920; AAC17915.1; -;  
DR EMBL, AF065921; AAC17916.2; -;  
DR EMBL, AF065922; AAC17917.2; -;  
DR EMBL, AF065923; AAC17918.2; -;  
DR PIR, A43271; A43271.  
DR PIR, A42166; A42166.  
DR PDB, 1NOC; 18-NOV-98.  
DR PDB, 1NOS; 18-NOV-98.  
DR PDB, 2NOS; 18-NOV-98.  
DR PDB, 1NOD; 23-MAR-99.  
DR PDB, 2NOD; 23-MAR-99.  
DR PDB, 3NOD; 23-MAR-99.  
DR PDB, 1DF1; 08-DEC-99.  
DR PDB, 1DMV; 04-FEB-00.  
DR PDB, 1DMW; 06-FEB-00.  
DR PDB, 1DMX; 06-FEB-00.  
DR PDB, 1DOW; 15-DEC-99.  
DR PDB, 1DD7; 29-MAR-00.  
DR MGD, MGI:97361; Nos2.  
DR InterPro: IPR003097; FAD\_binding.  
DR InterPro: IPR001709; FPN\_cyt\_redcse.  
DR InterPro: IPR001094; Flavodoxin\_like.  
DR InterPro: IPR001226; Flavodoxin.  
DR InterPro: IPR004030; NO\_synthase.  
DR InterPro: PF00175; NAD\_binding\_1.  
DR Pfam: PF00258; flavodoxin; 1.  
DR Pfam: PF00667; FAD\_binding; 1.  
DR Pfam: PF02898; NO\_synthase; 1.  
DR PRINTS: PR00369; FLAVODOXIN.  
DR PRINTS: PR00371; FPNCR.  
DR PROSITE, PS60001; NOS; 1.



KM Oxidoreductase; NADP; FAD; FMN; Calmodulin-binding; Heme;  
 KM Zinc; Metal-binding; Multigene family; 3D-structure.  
 FT BINDING 194 194 HEME.  
 FT DOMAIN 503 523 CALMODULIN-BINDING (POTENTIAL).  
 FT NP\_BIND 617 648 FMN (PYRIMIDINE PART) (BY SIMILARITY).  
 FT NP\_BIND 761 772 FAD (ADP PART) (BY SIMILARITY).  
 FT NP\_BIND 897 907 FAD (FLAVIN PART) (BY SIMILARITY).  
 FT NP\_BIND 972 990 NADP (RIBOSE PART) (BY SIMILARITY).  
 FT NP\_BIND 1070 1085 NADP (ADP PART) (BY SIMILARITY).  
 FT METAL 104 104 ZINC.  
 FT METAL 109 109 ZINC.  
 FT VARIANT 211 211 C -> R (IN STRAIN MOD/LTU).  
 FT VARIANT 967 967 P -> L (IN STRAIN SJL/J).  
 FT VARIANT 968 968 S -> F (IN STRAIN BALB/CBYJ).  
 FT CONFLICT 19 19 K -> T (IN REF. 4).  
 FT CONFLICT 191 191 A -> V (IN REF. 2).  
 FT CONFLICT 844 844 A -> G (IN REF. 2).  
 SQ SEQUENCE 1144 AA; 130574 MW; 0735BE676113457F CRC64;  
 Query Match 52.6%; Score 41; DB 1; Length 1144;  
 Best Local Similarity 50.0%; Pred. No. 50;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 Oy 1 AMEGPALYEDPPD 14  
 Db 305 ADCQPEVEFEIPPD 318  
 RESULT 14  
 NOS2\_RAT  
 ID NOS2\_RAT STANDARD; PRT: 1147 AA.  
 AC 006518; P97774; 035765; 035766; 064558; 064005; 063267;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Nitric oxide synthase, inducible (EC 1.14.13.39) (NOS, type II)  
 DE (Inducible NOS) (INOS).  
 GN NOS2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Vascular smooth muscle;  
 RX MEDLINE-93191721; PubMed7680561;  
 RA Nunokawa Y., Ishida N., Tanaka S.;  
 RT "Cloning of inducible nitric oxide synthase in rat vascular smooth  
 RT muscle cells.";  
 RL Biochem. Biophys. Res. Commun. 191:89-94(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Mistar; TISSUE-Pancreatic islets;  
 RX MEDLINE-95309542; PubMed7540573;  
 RA Karlsson A.E., Andersen H.U., Vissing H., Larsen P.M., Fey S.J.,  
 RA Cuartero B.G., Madsen O.D., Petersen J.S., Mortensen S.B.,  
 RA Mandrup-Poulsen T., Boel E., Nerup J.;  
 RT "Cloning and expression of cytokine-inducible nitric oxide synthase  
 RT cDNA from rat islets of Langerhans.";  
 RL Diabetes 44:753-758(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Sprague-Dawley; TISSUE-Astrocytes;  
 RX MEDLINE-94331594; PubMed7513765;  
 RA Galea E., Reis D.J., Feinstein D.L.;  
 RT "Cloning and expression of inducible nitric oxide synthase from rat  
 RT astrocytes.";  
 RL J. Neurosci. Res. 37:406-414(1994).  
 RN [4]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE-Liver;  
 RX MEDLINE-94039059; PubMed-7693462;  
 RA Adachi H., Iida S., Oguchi S., Ohshima H., Suzuki H., Nagasaki K.,  
 RA Kawasaki H., Sugimura T., Esumi H.;  
 RT "Molecular cloning of a cDNA encoding an inducible  
 RT calmodulin-dependent nitric-oxide synthase from rat liver and its  
 RT expression in COS 1 cells.";  
 RL Eur. J. Biochem. 217:37-43(1993).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Sprague-Dawley; TISSUE-Hepatocytes;  
 RX MEDLINE-93221515; PubMed-7682072;  
 RA Wood E.R., Berger H., Jr., Sherman P.A., Lapetina E.G.;  
 RT "Hepatocytes and macrophages express an identical cytokine inducible  
 RT nitric oxide synthase gene.";  
 RL Biochem. Biophys. Res. Commun. 191:767-774(1993).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Sprague-Dawley; TISSUE-Aorta;  
 RX MEDLINE-94325351; PubMed-7519448;  
 RA Geng Y.J., Almqvist M., Hansson G.K.;  
 RT "cDNA cloning and expression of inducible nitric oxide synthase from  
 RT rat vascular smooth muscle cells.";  
 RL Biochim. Biophys. Acta 1218:421-424(1994).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RA Kosuga K., Yui Y., Hattori R., Sase K., Ezawa H., Aoyama T.,  
 RA Inoue R., Sasayama S.;  
 RT "Cloning of an inducible nitric oxide synthase from rat  
 RT polymorphonuclear neutrophils.";  
 RL Endothelium 2:217-221(1994).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-97070590; PubMed-8913516;  
 RA Tsutsunishita Y., Kawai Y., Takahara H., Onda T., Miyoshi J.,  
 RA Futaki S., Niva M.;  
 RT "Sequence analysis of inducible nitric oxide synthase in rat kidney,  
 RT lung, and uterus.";  
 RL Biol. Pharm. Bull. 19:1374-1376(1996).  
 RN [9]  
 RP SEQUENCE OF 426-788 FROM N.A.  
 RC STRAIN-Dahl/Rapp salt sensitive strain; TISSUE-Vascular smooth muscle;  
 RX MEDLINE-98195092; PubMed-9535415;  
 RA Chen P.Y., Gladish R.D., Sanders P.M.;  
 RT "Vascular smooth muscle nitric oxide synthase anomalies in Dahl/Rapp  
 RT salt-sensitive rats.";  
 RL Hypertension 31:918-924(1998).  
 RN [10]  
 RP SEQUENCE OF 509-740 FROM N.A.  
 RC STRAIN-Mistar; TISSUE-Renal glomerulus;  
 RA Saura M., Zaragoza C., Martinez-Dalmau R., Perez-Sala D., Lamas S.;  
 RT "Advances in the studies of NO synthesis regulation in mesangial  
 RT cells.";  
 RL Nephrologia 16:35-39(1996).  
 RN [11]  
 RP SEQUENCE OF 479-655 FROM N.A.  
 RC STRAIN-Sprague-Dawley; TISSUE-Renal glomerulus;  
 RX MEDLINE-94276509; PubMed-7516453;  
 RA Morrissey J.J., McCracken R., Kaneto H., Vehaskari M., Montani D.,  
 RA Klahr S.;  
 RT "Location of an inducible nitric oxide synthase mRNA in the normal  
 RT kidney.";  
 RL Kidney Int. 45:998-1005(1994).  
 RN [12]  
 RP SEQUENCE OF 420-479 FROM N.A.  
 RC TISSUE-Myocardium;  
 RX MEDLINE-9420479 FROM N.A.  
 RA Mitchell T., Balligand J.-L.;  
 RT "Isolation and characterization of INOS from rat cardiocytes.";  
 RL Submitted (SEP-1994) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: PRODUCES NITRIC OXIDE (NO) WHICH IS A MESSENGER MOLECULE  
 CC WITH DIVERSE FUNCTIONS THROUGHOUT THE BODY.  
 CC -1- CATALYTIC ACTIVITY: L-arginine + N NADPH + M O(2) -> citrulline +  
 CC nitric oxide + N NADP(+).  
 CC -1- COFACTOR: HEME. BINDS ONE MOLE EACH OF FAD AND FMN. ALSO REQUIRES  
 CC TETRAHYDROBIOTERIN (BH4) WHICH MAY STABILIZE THE DIMERIC FORM OF  
 CC THE ENZYME.

FT	CONFLICT	412	412	V -> A (IN REF. 3).
FT	CONFLICT	477	477	M -> I (IN REF. 12).
FT	CONFLICT	513	513	T -> R (IN REF. 10).
FT	CONFLICT	515	515	L -> W (IN REF. 11).
FT	CONFLICT	545	545	G -> R (IN REF. 11).
FT	CONFLICT	551	551	A -> R (IN REF. 10).
FT	CONFLICT	556	556	A -> S (IN REF. 11).
FT	CONFLICT	564	564	T -> N (IN REF. 11).
FT	CONFLICT	570	570	E -> D (IN REF. 11).
FT	CONFLICT	583	583	L -> P (IN REF. 5 AND 9).
FT	CONFLICT	591	591	G -> A (IN REF. 11).
Query Match				
Best Local Similarity	52.6%;	Score 41;	DB 1;	Length 1147;
Matches	7;	Conservative	3;	Mismatches
			4;	Indels
				Gaps
OY	1	AAGGPGALYEDPPD	14	
Db	308	AGGDPEVEFEIIPD	321	
RESULT 15				
ID	NS2D_HUMAN	STANDARD:	PRT:	1147 AA.
AC	O60591; O60604;			
DT	30-MAY-2000 (Ref., 39, Created)			
DT	30-MAY-2000 (Ref., 39, Last sequence update)			
DT	16-OCT-2001 (Ref., 40, Last annotation update)			
DE	Nitric oxide synthase, inducible IID (EC 1.14.13.39) (NOS, type II D) (inducible NOS) (iNOS).			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Skeletal muscle, and Heart muscle;			
RX	MEDLINE=9906690; PubMed=9851365;			
RA	Adams V., Kriebles S., Jiang H., Yu J., Rahmel A., Giesen S.,			
RA	Schuler G., Hambrecht R.;			
RT	"Complete coding sequence of inducible nitric oxide synthase from			
RT	human heart and skeletal muscle of patients with chronic heart			
RT	failure".			
RL	Nitric oxide 2:242-249(1998).			
CC	-I- FUNCTION: PRODUCES NITRIC OXIDE (NO) WHICH IS A MESSENGER MOLECULE WITH DIVERSE FUNCTIONS THROUGHOUT THE BODY. IN MACROPHAGES, NO MEDIATES TUMORICIDAL AND BACTERICIDAL ACTIONS. MAY BE INVOLVED IN THE PATHOGENESIS OF DILATED CARDIOMYOPATHY OR OF EXERCISE INTOLERANCE OBSERVED IN PATIENTS WITH CHRONIC HEART FAILURE.			
CC	-I- CATALYTIC ACTIVITY: L-arginine + NADPH + M O(2) = citrulline + nitric oxide + N NADP(+).			
CC	-I- COFACTOR:HEME. BINDS ONE MOLE EACH OF PAD AND FMN. ALSO REQUIRES TETRAHYDROBIOPTERIN (BH4) WHICH MAY STABILIZE THE DIMERIC FORM OF THE ENZYME (BY SIMILARITY).			
CC	-I- ENZYME REGULATION: REGULATED BY CALCIUM/CALMODULIN (BY SIMILARITY).			
CC	-I- SUBUNIT: HOMODIMER (BY SIMILARITY).			
CC	-I- TISSUE SPECIFICITY: EXPRESSED IN THE HEART AND SKELETAL MUSCLE DURING CHRONIC HEART FAILURE, BUT NOT IN HEALTHY INDIVIDUALS.			
CC	-I- SIMILARITY: BELONGS TO THE NOS FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; AF049656; AAC83553.1; -			
DR	EMBL; AF051164; AAC83554.1; -			
DR	HSSP; P29477; INOC.			
DR	InterPro; IPR003097; FAD_binding.			

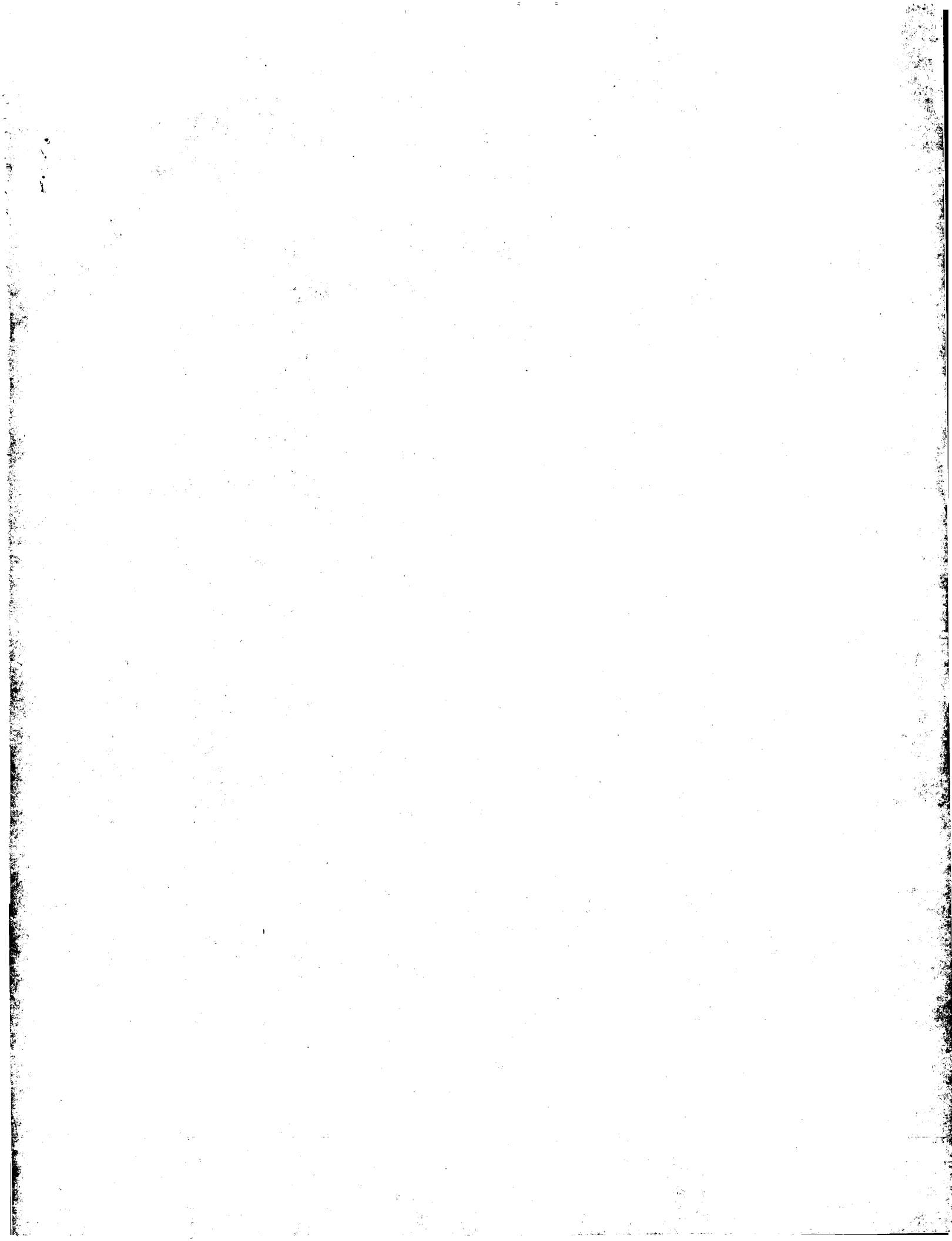
```

DR InterPro: IPR001709; FPN_cyl_redctase.
DR InterPro: IPR001094; Flavodoxin_like.
DR InterPro: IPR001226; Flavodoxin.
DR InterPro: IPR004030; NO_synthase.
DR InterPro: IPR001433; Oxired_FAD/NAD(P).
DR pfam: PF00175; NAD_binding; 1.
DR pfam: PF00258; Flavodoxin; 1.
DR pfam: PF00667; FAD_binding; 1.
DR pfam: PF02898; NO_synthase; 1.
DR PRINTS: PR00369; FLAVODOXIN.
DR PRINTS: PR00371; FPNCR.
DR PROSITE: PS60001; NOS; 1.
KW Oxidoreductase; NADP; FAD; FMN; Calcium-binding; Calmodulin-binding;
Heme; Zinc; Metal-binding; Multigene family.
FT BINDING 197 197 HEME (BY SIMILARITY).
FT DOMAIN 306 326 CALMODULIN-BINDING (POTENTIAL).
FT NP_BIND 620 651 FMN (PYRIMIDINE PART) (BY SIMILARITY).
FT NP_BIND 764 775 FAD (ADP PART) (BY SIMILARITY).
FT NP_BIND 900 910 FAD (FLAVIN PART) (BY SIMILARITY).
FT NP_BIND 975 993 NADP (RIBOSE PART) (BY SIMILARITY).
FT NP_BIND 1073 1088 NADP (ADP PART) (BY SIMILARITY).
FT METAL 107 107 ZINC (BY SIMILARITY).
FT METAL 112 112 ZINC (BY SIMILARITY).
FT CONFLICT 248 248 D -> S (IN REF. 1; AAC83554).
FT CONFLICT 271 271 G -> A (IN REF. 1; AAC83554).
FT CONFLICT 399 399 P -> Q (IN REF. 1; AAC83554).
FT CONFLICT 640 640 A -> T (IN REF. 1; AAC83554).
FT CONFLICT 731 731 H -> N (IN REF. 1; AAC83554).
FT CONFLICT 937 937 TL -> NF (IN REF. 1; AAC83554).
FT CONFLICT 1008 1009 E -> K (IN REF. 1; AAC83554).
FT CONFLICT 1024 1024 I -> L (IN REF. 1; AAC83554).
FT CONFLICT 1076 1076 F -> V (IN REF. 1; AAC83554).
FT CONFLICT 1129 1129
SQ SEQUENCE 1147 AA; 130528 MW; FF7E4C7ABA76D820 CRC64;

Query Match 52.6%; Score 41; DB 1; Length 1147;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAGEGPAlyEDPPD 14
DB 308 AHGDDPEVEIIPP 321

```

Search completed: January 30, 2003, 14:37:56  
 Job time : 13 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2003, 14:37:03 : Search time 14 Seconds  
(without alignments)  
29.423 Million cell updates/sec

Title: US-10-001-426-2  
Perfect score: 78  
Sequence: 1 AAGSGPALYEDPPD 14

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/6C.COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/6D.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	14	2	US-08-795-927-2
2	78	100.0	318	2	US-08-872-719-2
3	78	100.0	318	3	US-08-957-302A-12
4	78	100.0	318	4	US-09-336-890-2
5	78	100.0	318	4	US-09-542-403-12
6	78	100.0	318	4	US-09-668-499-2
7	78	100.0	319	4	US-09-668-499-2
8	78	100.0	524	3	US-08-957-302A-2
9	78	100.0	524	4	US-09-542-403-2
10	78	100.0	1146	4	US-09-126-109-12
11	78	100.0	1153	1	US-08-314-917-2
12	78	100.0	1153	1	US-08-265-046-2
13	78	100.0	1153	1	US-08-265-046-2
14	78	100.0	1153	5	PCT-US93-11401-2
15	78	100.0	1153	5	PCT-US93-07849-2
16	78	100.0	1182	4	US-09-287-354-6
17	78	100.0	1207	4	US-09-287-354-5
18	78	100.0	1144	1	US-08-147-812-5
19	78	100.0	1144	2	US-08-319-866-12
20	78	100.0	1144	4	US-09-123-708-2
21	78	100.0	1144	4	US-09-123-624-2
22	78	100.0	1189	4	US-09-287-354-4
23	78	100.0	717	4	US-09-307-143-2
24	78	100.0	855	4	US-08-890-865A-10
25	78	100.0	933	4	US-08-764-870-14
26	78	100.0	933	4	US-08-980-115-14
27	78	100.0	1447	4	US-09-041-886-25

28	38	48.7	1447	5	PCT-US94-05277-2	Sequence 2, Appl1
29	37	47.4	86	2	US-08-343-443B-7	Sequence 7, Appl1
30	37	47.4	355	4	US-08-630-915A-192	Sequence 192, App
31	37	47.4	656	2	US-08-343-443B-2	Sequence 2, Appl1
32	37	47.4	656	4	US-09-214-564A-4	Sequence 4, Appl1
33	37	47.4	707	5	US-08-528-122-18	Sequence 18, Appl1
34	37	47.4	707	5	PCT-US95-11720-18	Sequence 62, Appl
35	37	47.4	713	1	US-08-188-228-62	Sequence 56, Appl
36	37	47.4	713	1	US-08-332-643-56	Sequence 62, Appl
37	37	47.4	713	1	US-08-332-643-56	Sequence 62, Appl
38	37	47.4	750	1	US-08-325-553-2	Sequence 2, Appl1
39	37	47.4	750	2	US-08-394-152A-2	Sequence 2, Appl1
40	37	47.4	750	4	US-09-044-668-2	Sequence 6, Appl1
41	37	47.4	1336	5	US-08-551-356-6	Sequence 6, Appl1
42	37	47.4	1336	5	PCT-US93-12687-6	Sequence 2, Appl1
43	37	47.4	2446	2	US-08-551-356-2	Sequence 2, Appl1
44	37	47.4	2446	5	PCT-US93-12687-2	Sequence 328, App
45	36	46.2	71	4	US-09-188-930-328	

## ALIGNMENTS

RESULT 1  
US-08-795-927-2  
Sequence 2, Application US/08795927  
Patent No. 5986072  
GENERAL INFORMATION:  
APPLICANT: HANADA, Hiroshi  
APPLICANT: KAWAGUCHI, Haruma  
TITLE OF INVENTION: DRUG-IMMOBILIZED PARTICLES AND A  
TITLE OF INVENTION: PROCESS OF PURIFYING PROTEINS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTEST Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/795,927  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Corless, Peter F  
REGISTRATION NUMBER: 33,860  
REFERENCE/DOCKET NUMBER: 47115  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
US-08-795-927-2

Query Match 100.0%; Score 78; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 9.8e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAGEGPALEYDDPD 14  
Db 1 AAGEGPALEYDDPD 14

## RESULT 2

US-08-872-719-2  
Sequence 2, Application US/08872719

Patent No. 5919643

GENERAL INFORMATION:

APPLICANT: Kelley, Mark R.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE USE OF APURINIC/APYRIMIDINIC

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: ARNOLD, WHITE & DURKEE

STREET: P.O. BOX 4433

CITY: HOUSTON

STATE: TEXAS

COUNTRY: USA

ZIP: 77057-2198

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/872,719

FILING DATE: CONCURRENTLY HEREWITH

CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:

NAME: Highlander, Steven L.

REGISTRATION NUMBER: 37,642

REFERENCE/DOCKET NUMBER: INDY:012P21

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (713) 789-2679

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 318 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-872-719-2

Query Match 100.0%; Score 78; DB 2; Length 318;

Best Local Similarity 100.0%; Pred. No. 2.7e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAGEGPALEYDDPD 14  
Db 37 AAGEGPALEYDDPD 50

RESULT 3  
US-08-957-302A-12  
Sequence 12, Application US/08957302A

Patent No. 6046036

GENERAL INFORMATION:

APPLICANT: Kelley, Mark

TITLE OF INVENTION: DNA Sequences Encoding Fusions of DNA

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: ARNOLD, WHITE & DURKEE

STREET: P.O. BOX 4433

CITY: HOUSTON

STATE: TX

COUNTRY: US

ZIP: 77210-4433

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/957,302A

FILING DATE: CONCURRENTLY HEREWITH

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Highlander, Steven L.

REGISTRATION NUMBER: 37,642

REFERENCE/DOCKET NUMBER: INDY:005

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (713) 789-2679

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 318 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-957-302A-12

Query Match 100.0%; Score 78; DB 3; Length 318;

Best Local Similarity 100.0%; Pred. No. 2.7e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAGEGPALEYDDPD 14  
Db 37 AAGEGPALEYDDPD 50

RESULT 4  
US-09-336-890-2  
Sequence 2, Application US/09336890

Patent No. 6190661

GENERAL INFORMATION:

APPLICANT: Kelley, Mark R.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE USE OF

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: ARNOLD, WHITE & DURKEE

STREET: P.O. BOX 4433

CITY: HOUSTON

STATE: TEXAS

COUNTRY: USA

ZIP: 77057-2198

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/336, 890

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/872,719

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Highlander, Steven L.

REGISTRATION NUMBER: 37,642

REFERENCE/DOCKET NUMBER: INDY:012P21

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (713) 789-2679

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 318 amino acids

TYPE: amino acid

```

: STRANDEDNESS:
: TOPOLOGY: linear
US-09-336-890-2
Query Match 100.0%; Score 78; DB 4; Length 318;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAGEGPAlyEDPPD 14
Db 37 AAGEGPAlyEDPPD 50

RESULT 5
US-09-542-403-12
: Sequence 12, Application US/09542403
: Patent No. 6252048
: GENERAL INFORMATION:
: APPLICANT: Kelley, Mark
: ATTORNEY/AGENT INFORMATION:
: TITLE OF INVENTION: DNA Sequences Encoding Fusions of DNA
: TITLE OF INVENTION: Repair Proteins and Uses Thereof
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ARNOLD, WHITE & DURKEE
: STREET: P.O. BOX 4433
: CITY: HOUSTON
: STATE: TX
: COUNTRY: US
: ZIP: 77210-4433
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/542,403
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/957,302
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Highlander, Steven L.
: REGISTRATION NUMBER: 37,642
: REFERENCE/DOCKET NUMBER: INDY.005
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (512) 418-3000
: TELEFAX: (713) 789-2679
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 318 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
US-09-542-403-12
Query Match 100.0%; Score 78; DB 4; Length 318;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAGEGPAlyEDPPD 14
Db 37 AAGEGPAlyEDPPD 50

RESULT 6
US-09-668-499-2
: Sequence 2, Application US/09668499
: Patent No. 6406917
: GENERAL INFORMATION:
: APPLICANT: Kelley, Mark R.
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE USE OF APURINIC/ADYRIMIDINIC
```

```

: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ARNOLD, WHITE & DURKEE
: STREET: P.O. BOX 4433
: CITY: HOUSTON
: STATE: TEXAS
: COUNTRY: USA
: ZIP: 77057-2198
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/668,499
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/872,719
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Highlander, Steven L.
: REGISTRATION NUMBER: 37,642
: REFERENCE/DOCKET NUMBER: INDY:012P21
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (512) 418-3000
: TELEFAX: (713) 789-2679
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 318 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
US-09-668-499-2
Query Match 100.0%; Score 78; DB 4; Length 318;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAGEGPAlyEDPPD 14
Db 37 AAGEGPAlyEDPPD 50

RESULT 7
US-08-795-927-4
: Sequence 4, Application US/08795927
: Patent No. 5986072
: GENERAL INFORMATION:
: APPLICANT: HANDA, Hitoshi
: ATTORNEY/AGENT INFORMATION:
: TITLE OF INVENTION: DRUG-IMMOBILIZED PARTICLES AND A
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BronsteIn, Roberts & Cushman, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/795,927
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
```

ATTORNEY/AGENT INFORMATION:  
NAME: Corless, Peter F  
REGISTRATION NUMBER: 33,860  
REFERENCE/DOCKET NUMBER: 47115  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 319 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
US-08-795-927-4

Query Match 100.0%; Score 78; DB 2; Length 319;  
Best Local Similarity 100.0%; Pred. No. 2.7e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGEGPALYEDPPD 14  
Db 37 AAGEGPALYEDPPD 50

## RESULT 8

US-08-957-302A-2  
Sequence 2, Application US/08957302A

GENERAL INFORMATION:  
APPLICANT: Kelley, Mark  
TITLE OF INVENTION: DNA Sequences Encoding Fusions of DNA  
TITLE OF INVENTION: Repair Proteins and Uses Thereof  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARNOLD, WHITE & DURKEE  
STREET: P.O. BOX 4433  
CITY: HOUSTON  
STATE: TX  
COUNTRY: US  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/957,302A  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: INDY:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 524 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-957-302A-2

Query Match 100.0%; Score 78; DB 3; Length 524;  
Best Local Similarity 100.0%; Pred. No. 4.6e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AAGEGPALYEDPPD 14  
Db 243 AAGEGPALYEDPPD 256

## RESULT 9

US-09-542-403-2  
Sequence 2, Application US/09542403

GENERAL INFORMATION:  
APPLICANT: Kelley, Mark  
TITLE OF INVENTION: DNA Sequences Encoding Fusions of DNA  
TITLE OF INVENTION: Repair Proteins and Uses Thereof  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARNOLD, WHITE & DURKEE  
STREET: P.O. BOX 4433  
CITY: HOUSTON  
STATE: TX  
COUNTRY: US  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/542,403  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/957,302

ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: INDY:005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 524 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-09-542-403-2

Query Match 100.0%; Score 78; DB 4; Length 524;  
Best Local Similarity 100.0%; Pred. No. 4.6e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGEGPALYEDPPD 14  
Db 243 AAGEGPALYEDPPD 256

## RESULT 10

US-09-126-109-12  
Sequence 12, Application US/09126109

GENERAL INFORMATION:  
APPLICANT: Thiipen, Anice  
APPLICANT: Hohmeier, Hans-Ewald  
APPLICANT: Newgard, Christopher B.  
APPLICANT: Unger, Roger H.  
APPLICANT: Shimabukuro, Michio  
APPLICANT: Chen, Guaxun  
APPLICANT: Rhodes, Christopher J.  
APPLICANT: Hugl, Stigrun R.



APPLICANT: Cousin, Sharon  
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING  
TO NO-MEDICATED CYTOTOXICITY  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/126,109  
FILING DATE: 30-JUL-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/055,092  
FILING DATE: 30-JUL-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US Unknown  
FILING DATE: 03-MAR-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: McMillian, Nabeela R.  
REGISTRATION NUMBER: P-43,363  
REFERENCE/DOCKET NUMBER: UTS:560  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1146 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-126-109-12

Query Match 53.8%; Score 42; DB 4; Length 1146;  
Best Local Similarity 57.1%; Pred. No. 80;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 AAGEGPALEDPPD 14  
| | | | |  
Db 311 ANGRDPELEIPD 324

RESULT 11  
US-08-314-917-2  
Sequence 2, Application US/08314917  
Patent No. 5468630  
GENERAL INFORMATION:  
APPLICANT: Billiar, Timothy R.  
APPLICANT: Nusser, Andreas K.  
APPLICANT: Geller, David A.  
APPLICANT: Simmons, Richard L.  
TITLE OF INVENTION: CDNA Clone for Human Inducible Nitric  
Oxide Synthase And Process for Preparing Same  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold B. Silverman  
STREET: Eckert Seamans Cherin & Melliott  
CITY: Pittsburgh  
STATE: PA  
COUNTRY: USA  
ZIP: 15219  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/314,917  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/981,344  
FILING DATE: 25-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Silverman, Arnold B.  
REGISTRATION NUMBER: 22,614  
REFERENCE/DOCKET NUMBER: 116972  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (412) 566-6000  
TELEFAX: (412) 566-6099  
TELEX: 866172  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1153 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-314-917-2

Query Match 53.8%; Score 42; DB 1; Length 1153;  
Best Local Similarity 57.1%; Pred. No. 80;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 AAGEGPALEDPPD 14  
| | | | |  
Db 311 ANGRDPELEIPD 324

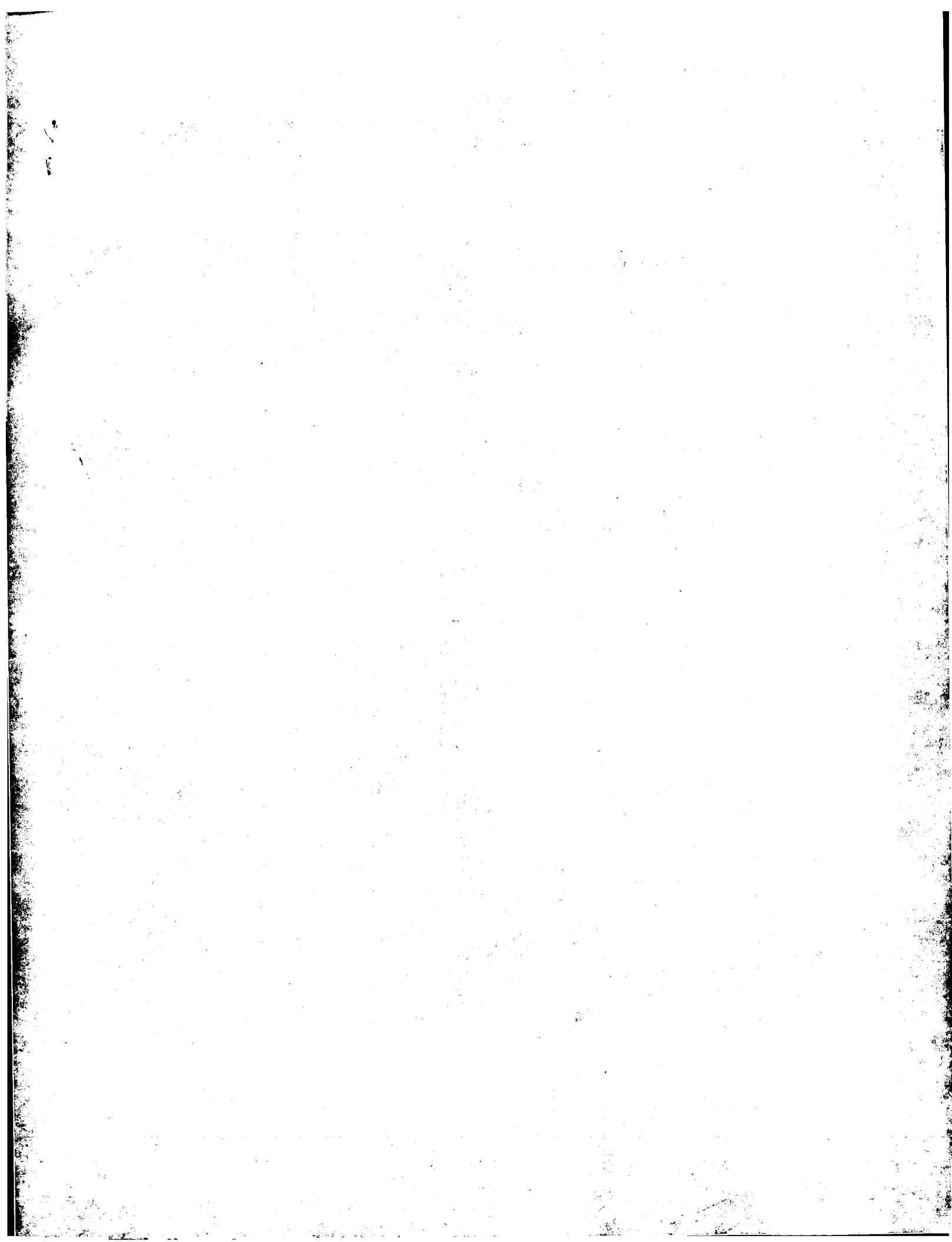
RESULT 12  
US-08-265-046-2  
Sequence 2, Application US/08265046  
Patent No. 5658565  
GENERAL INFORMATION:  
APPLICANT: Timothy R. Billiar  
APPLICANT: Edith Tzeng  
APPLICANT: Andreas K. Nusser  
APPLICANT: David A. Geller  
APPLICANT: Richard L. Simmons  
TITLE OF INVENTION: Inducible Nitric Oxide Synthase  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lewis F. Gould, Jr.  
STREET: Eckert Seamans Cherin & Melliott  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/265,046  
FILING DATE: 24-JUN-1994  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Gould, Lewis F. Jr.  
REGISTRATION NUMBER: 25,057  
REFERENCE/DOCKET NUMBER: 119130  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 575-6020  
TELEFAX: (215) 575-6015  
TELEX:



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/07849  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: GOULD, Lewis F., Jr.  
REGISTRATION NUMBER: 25,057  
REFERENCE/DOCKET NUMBER: 119130-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 575-6020  
TELEFAX: (215) 575-6015  
TELEX:  
INFORMATION FOR SEQ. ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1153 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-07849-2

Query Match 53.8%; Score 42; DB 5; Length 1153;  
Best Local Similarity 57.1%; Pred. No. 80;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
QY 1 AGEGPALYEDPPD 14  
| | | | |  
Db 311 ANGRDPELFEIIPP 324

Search completed: January 30, 2003, 14:39:15  
Job time: 16 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2003, 14:38:39 : Search time 10 Seconds  
(without alignments)  
28.250 Million cell updates/sec

Title: US-10-001-426-2

Perfect score: 78

Sequence: 1 AAGEGALYEDPPD 14

Scoring table: BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

Published Applications-AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	14	9	US-10-001-426-2
2	78	100.0	14	9	US-10-003-014-2
3	78	100.0	319	9	US-10-001-426-4
4	78	100.0	319	9	US-10-003-014-4
5	78	100.0	345	10	US-09-925-300-1188
6	44	56.4	68	10	US-09-925-297-551
7	42	53.8	1153	10	US-09-925-056-2
8	41	52.6	1144	9	US-09-870-759-124
9	39.5	50.6	561	9	US-09-822-364A-1
10	39.5	50.6	561	9	US-09-254-590-1
11	39.5	50.6	561	9	US-10-115-695-1
12	39	50.0	201	10	US-09-925-300-1432
13	39	50.0	563	9	US-09-738-626-6643
14	39	50.0	7968	9	US-10-077-130-5
15	38	48.7	114	10	US-09-994-365-3
16	38	48.7	136	10	US-09-994-365-2
17	38	48.7	456	10	US-09-815-242-10400
18	38	48.7	456	10	US-09-815-242-11073
19	38	48.7	456	10	US-09-815-242-14064

20	38	48.7	458	10	US-09-815-242-11759	Sequence 11759, A
21	37	47.4	15	9	US-09-942-117-14	Sequence 14, App11
22	37	47.4	91	9	US-09-942-117-4	Sequence 4, App11
23	37	47.4	171	10	US-09-925-301-1078	Sequence 1078, Ap
24	37	47.4	355	10	US-09-879-957-192	Sequence 192, App
25	37	47.4	368	10	US-09-925-299-888	Sequence 888, App
26	37	47.4	378	10	US-09-864-761-43251	Sequence 43251, A
27	37	47.4	448	10	US-09-939-521-7	Sequence 7, App11
28	37	47.4	733	10	US-09-862-027-42	Sequence 42, App11
29	37	47.4	750	9	US-09-978-295A-618	Sequence 618, App
30	37	47.4	750	9	US-09-978-697-618	Sequence 618, App
31	37	47.4	750	9	US-09-978-192A-618	Sequence 618, App
32	37	47.4	750	9	US-09-895-793-944	Sequence 944, App
33	37	47.4	750	9	US-09-999-832A-618	Sequence 618, App
34	37	47.4	750	9	US-09-978-189-618	Sequence 618, App
35	37	47.4	750	9	US-10-174-590-104	Sequence 104, App
36	37	47.4	750	9	US-10-176-758-104	Sequence 104, App
37	37	47.4	750	9	US-10-175-737-104	Sequence 104, App
38	37	47.4	750	10	US-09-823-827-944	Sequence 944, App
39	37	47.4	750	12	US-10-052-586-104	Sequence 104, App
40	37	47.4	3472	9	US-10-027-806-4	Sequence 4, App11
41	37	47.4	3472	9	US-10-034-623-4	Sequence 4, App11
42	36	46.2	71	9	US-09-981-876-239	Sequence 239, App
43	36	46.2	71	9	US-09-981-876-240	Sequence 240, App
44	36	46.2	88	10	US-09-864-761-42871	Sequence 42871, A
45	36	46.2	204	10	US-09-922-378-3	Sequence 3, App11

#### ALIGNMENTS

RESULT 1  
US-10-001-426-2  
; Sequence 2, Application US/10001426  
; Patent No. US20020160472A1  
GENERAL INFORMATION:  
APPLICANT: HANDA, Hiroshi  
TITLE OF INVENTION: DRUG-IMMOBILIZED PARTICLES AND A  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dike, Bronsteijn, Roberts & Cushman, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/001,426  
FILING DATE: 02-NO. US20020160472A1-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/440,597  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Corleas, Peter F  
REGISTRATION NUMBER: 33,860  
REFERENCE/DOCKET NUMBER: 47115  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO. 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single



```

; KAWAGUCHI, Haruma
; TITLE OF INVENTION: DRUG-IMMOBILIZED PARTICLES AND A
; PROCESS OF PURIFYING PROTEINS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, Bronsteijn, Roberts & Cushman, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/003,014
; FILING DATE: 02-NO. US20020169293A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/440,597
; FILING DATE: 15-NO. US20020169293A1-1999
; APPLICATION NUMBER: 08/795,927
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Corleese, Peter F
; REGISTRATION NUMBER: 33,860
; REFERENCE/DOCKET NUMBER: 47115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 319 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-003-014-4
Query Match 100.0%; Score 78; DB 9; Length 319;
Best Local Similarity 100.0%; Pred. No. 5.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGECPALYEDPPD 14
DB 37 AAGEGPALYEDPPD 50
RESULT 5
US-09-925-300-1188
; Sequence 1188, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
```

```

; SEQ ID NO 1188
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (175)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1188
Query Match 100.0%; Score 78; DB 10; Length 345;
Best Local Similarity 100.0%; Pred. No. 6.3e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGECPALYEDPPD 14
DB 64 AAGEGPALYEDPPD 77
RESULT 6
US-09-925-297-551
; Sequence 551, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 551
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-551
Query Match 56.4%; Score 44; DB 10; Length 68;
Best Local Similarity 61.5%; Pred. No. 1.9;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAGECPALYEDPP 13
DB 39 AAGECPALYEDPP 51
RESULT 7
US-09-992-056-2
; Sequence 2, Application US/09992056
; Patent No. US20020061862A1
; GENERAL INFORMATION:
; APPLICANT: Ballier, Timothy R.
; APPLICANT: Tzend, Edith
; APPLICANT: Nussler, Andreas K.
; APPLICANT: Geller, David A.
; APPLICANT: Simmons, Richard K.
; TITLE OF INVENTION: Inducible Nitric Oxide Synthase Gene for Treatment of
; FILE REFERENCE: 213298
; CURRENT APPLICATION NUMBER: US/09/992,056
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 09/176,496
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 08/465,522
```

```
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: 08/314,917
; PRIOR FILING DATE: 1994-09-28
; PRIOR APPLICATION NUMBER: 07/981,344
; PRIOR FILING DATE: 1992-11-25
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 2
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Induced Human Hepatocyte RNA
US-09-922-056-2

Query Match          53.8%; Score 42; DB 10; Length 1153;
Best Local Similarity 57.1%; Pred. No. 73;
Matches      8; Conservative      1; Mismatches      5; Indels      0; Gaps      0;

QY      1 AAGEGPALEYEDPPD 14
DB      311 AANGDELEFEIPPD 324

RESULT 8
US-09-870-759-124
; Sequence 124, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 124
; LENGTH: 1144
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-870-759-124

Query Match          52.6%; Score 41; DB 9; Length 1144;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches      7; Conservative      3; Mismatches      4; Indels      0; Gaps      0;

QY      1 AAGEGPALEYEDPPD 14
DB      305 ADGDPEVEFEIPPD 318

RESULT 9
US-09-922-364A-1
; Sequence 1, Application US/09922364A
; Patent No. US2002015531A1
; GENERAL INFORMATION:
; APPLICANT: Adelman, John P.
; Maylie, James
; Bond, Chris T.
; Silvia, Christopher P.
; TITLE OF INVENTION: Small and Intermediate Conductance,
; Calcium-Activated Potassium Channels and Uses
; Thereof
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/922,364A
FILING DATE: 03-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/254,590
FILING DATE: 10-Mar-1999
APPLICATION NUMBER: US 60/026,451
FILING DATE: 11-SEP-1996
APPLICATION NUMBER: US 60/040,052
FILING DATE: 07-MAR-1997
APPLICATION NUMBER: US 60/045,233
FILING DATE: 17-APR-1997
APPLICATION NUMBER: WO PCT/US97/16033
FILING DATE: 10-SEP-1997
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 014210-000730US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 561 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: protein
LOCATION: 1..561
OTHER INFORMATION: /note="human small conductance,
calcium-activated potassium channel
protein 1 (hSK1)"
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-922-364A-1

Query Match          50.6%; Score 39.5; DB 9; Length 561;
Best Local Similarity 69.2%; Pred. No. 84;
Matches      9; Conservative      0; Mismatches      3; Indels      1; Gaps      1;

QY      3 GEGP-ALYEDPPD 14
DB      33 GSGPGLGRDPPD 45

RESULT 10
US-09-254-590-1
; Sequence 1, Application US/09254590
; Patent No. US20020165379A1
; GENERAL INFORMATION:
; APPLICANT: Adelman, John P.
; Maylie, James
; Bond, Chris T.
; Silvia, Christopher P.
; TITLE OF INVENTION: Small and Intermediate Conductance,
; Calcium-Activated Potassium Channels and Uses Thereof
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
```



SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/254,590  
FILING DATE: 10-Mar-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/026,451  
FILING DATE: 11-SEP-1996  
APPLICATION NUMBER: US 60/040,052  
FILING DATE: 07-MAR-1997  
APPLICATION NUMBER: US 60/045,233  
FILING DATE: 17-APR-1997  
APPLICATION NUMBER: WO PCT/US97/16033  
FILING DATE: 10-SEP-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 014210-000730US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 561 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..561  
OTHER INFORMATION: /note="human small conductance,  
calcium-activated potassium channel  
protein 1 (hsk1)"  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-254-590-1  
Query Match 50.6%; Score 39.5; DB 9; Length 561;  
Best Local Similarity 69.2%; Pred. No. 84;  
Matches 9; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
QY 3 GEGP-ALYEDPPD 14  
Db 33 GSGPGALGRDPPD 45  
RESULT 11  
US-10-115-695-1  
Sequence 1, Application US/10115695  
Publication No. US20020192757A1  
GENERAL INFORMATION:  
APPLICANT: Adelman, John P.  
Maylie, James  
Bond, Chris T.  
Silvia, Christopher P.  
TITLE OF INVENTION: Small and Intermediate Conductance,  
Calcium-Activated Potassium Channels and Uses  
Thereof  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/115,695

FILING DATE: 03-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/254,590  
FILING DATE: 10-Mar-1999  
APPLICATION NUMBER: US 60/026,451  
FILING DATE: 11-SEP-1996  
APPLICATION NUMBER: US 60/040,052  
FILING DATE: 07-MAR-1997  
APPLICATION NUMBER: US 60/045,233  
FILING DATE: 17-APR-1997  
APPLICATION NUMBER: WO PCT/US97/16033  
FILING DATE: 10-SEP-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 014210-000730US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 561 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..561  
OTHER INFORMATION: /note="human small conductance,  
calcium-activated potassium channel  
protein 1 (hsk1)"  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-115-695-1  
Query Match 50.6%; Score 39.5; DB 9; Length 561;  
Best Local Similarity 69.2%; Pred. No. 84;  
Matches 9; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
QY 3 GEGP-ALYEDPPD 14  
Db 33 GSGPGALGRDPPD 45  
RESULT 12  
US-09-925-300-1432  
Sequence 1432, Application US/09925300  
Patent No. US20020151681A1  
GENERAL INFORMATION:  
APPLICANT: Craig Rosen,  
Steve Ruben  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA101  
CURRENT APPLICATION NUMBER: US/09/925,300  
PRIOR FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05988  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 1690  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 1432  
LENGTH: 201  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (114)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (193)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE  
 ; LOCATION: (201)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 US-09-925-300-1432

Query Match	50.0%;	Score 39;	DB 10;	Length 201;
Best Local Similarity	72.7%;	Pred. No. 34;		
Matches	8;	Conservative	0;	Mismatches 3;
			Indels	0;
			Gaps	0;

```
QY      3 GEGPALYEDPP 13
        || || || ||
Db      70 GEEELKEDPP 80
```

```

RESULT 13
US-09-738-626-6643
: Sequence 6643, Application US/09738626
: Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAMA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MAKIKO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHITO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738, 626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: Patentln ver. 3.0
SEQ ID NO 6643
LENGTH: 563
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6643

```

Query Match	50.0%;	Score 39;	DB 9;	Length 563;
Best Local Similarity	53.3%;	Pred. No. 1e+02;		
Matches	8;	Conservative	2;	Mismatches 1;
				Indels 4;
				Gaps 1

```
QY      3 GEGP---ALYEDPP 13
          |::| |::| |
Db      319 GDGPRTDAIYEDTP 333
```

```

RESULT 14
US-10-077-130-5
; Sequence 5, Application US/10077130
; Patent No. US20020168742A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MP12001-047P1RCPl(M)
; CURRENT APPLICATION NUMBER: US/10/077,130
; PRIORITY FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5

```

```

; LENGTH: 7968
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-077-130-5

```

Query Match	50.0%;	Score 39;	DB 9;	Length 7968;
Best Local Similarity	57.1%;	Pred. No. 1.6e+03;		
Matches	8;	Mismatches	4;	Indels 0;
				Gaps 0;

```
QY      1  AAGEGPALYEDPPD 14
        : | | | | | | : | |
Db      5034 SADEGPALPEEPAD 5047
```

```

RESULT 15
US-09-994-365-3
: Sequence 3, Application US/09994365
: Patent No. US20020115148A1
: GENERAL INFORMATION:
: APPLICANT: Charmley, Patrick
: APPLICANT: Moss, Patrick
: APPLICANT: McEuen, Mark
: TITLE OF INVENTION: Compositions and Methods for Diagnosing or Treating Psoriasis
: FILE REFERENCE: CEC118109
: CURRENT APPLICATION NUMBER: US/09/994,365
: CURRENT FILING DATE: 2001-11-26
: PRIOR APPLICATION NUMBER: US 60/253,592
: PRIOR FILING DATE: 2000-11-28
: PRIOR APPLICATION NUMBER: US 60/256,839
: PRIOR FILING DATE: 2000-12-15
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO: 3
: LENGTH: 114
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-994-365-3

```

Query Match	48.7%	Score 38	DB 10	Length 114
Best Local Similarity	75.0%	Pred No. 27		
Matches 6	Conservative 1	Mismatches 0	Gaps 0	

```

QY      6 PALYEDPP 13
        | : | | |
Db      36 PLEFEDPP 43

```

Search completed: January 30, 2003, 14:42:23  
Job time : 11 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2003, 14:37:44 : Search time 139 Seconds

(Without alignments)  
64.937 Million cell updates/sec

Title: US-10-001-426-2

Perfect score: 78  
Sequence: 1 AAGEGPALEDPPD 14

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

Pending Patents\_AA\_Main:\*

1:	/cgn2_6/ptodata/1/paa/PCTRUS.COMB.pep.*
2:	/cgn2_6/ptodata/1/paa/US06.COMB.pep.*
3:	/cgn2_6/ptodata/1/paa/US07.COMB.pep.*
4:	/cgn2_6/ptodata/1/paa/US08.COMB.pep.*
5:	/cgn2_6/ptodata/1/paa/US081.COMB.pep.*
6:	/cgn2_6/ptodata/1/paa/US082.COMB.pep.*
7:	/cgn2_6/ptodata/1/paa/US083.COMB.pep.*
8:	/cgn2_6/ptodata/1/paa/US084.COMB.pep.*
9:	/cgn2_6/ptodata/1/paa/US085.COMB.pep.*
10:	/cgn2_6/ptodata/1/paa/US086.COMB.pep.*
11:	/cgn2_6/ptodata/1/paa/US087.COMB.pep.*
12:	/cgn2_6/ptodata/1/paa/US088.COMB.pep.*
13:	/cgn2_6/ptodata/1/paa/US089.COMB.pep.*
14:	/cgn2_6/ptodata/1/paa/US090.COMB.pep.*
15:	/cgn2_6/ptodata/1/paa/US091.COMB.pep.*
16:	/cgn2_6/ptodata/1/paa/US092.COMB.pep.*
17:	/cgn2_6/ptodata/1/paa/US093.COMB.pep.*
18:	/cgn2_6/ptodata/1/paa/US094.COMB.pep.*
19:	/cgn2_6/ptodata/1/paa/US095.COMB.pep.*
20:	/cgn2_6/ptodata/1/paa/US096.COMB.pep.*
21:	/cgn2_6/ptodata/1/paa/US097.COMB.pep.*
22:	/cgn2_6/ptodata/1/paa/US098.COMB.pep.*
23:	/cgn2_6/ptodata/1/paa/US099.COMB.pep.*
24:	/cgn2_6/ptodata/1/paa/US100.COMB.pep.*
25:	/cgn2_6/ptodata/1/paa/US101.COMB.pep.*
26:	/cgn2_6/ptodata/1/paa/US102.COMB.pep.*
27:	/cgn2_6/ptodata/1/paa/US60.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	78	100.0	14	US-09-440-597-2
2	78	100.0	14	US-10-001-426-2
3	78	100.0	14	US-10-003-014-2
4	78	100.0	69	US-10-029-386-28638
5	78	100.0	73	US-09-834-366-16628
6	78	100.0	73	US-60-197-873-16628

7	78	100.0	146	US-09-791-537-41381	Sequence 41381, A
8	78	100.0	170	PCT-US01-18569-4083	Sequence 4083, Ap
9	78	100.0	170	US-10-264-049-4083	Sequence 4083, Ap
10	78	100.0	183	US-09-760-482-73	Sequence 73, Appl
11	78	100.0	183	US-09-760-485-815	Sequence 815, Appl
12	78	100.0	183	US-10-153-088-73	Sequence 73, Appl
13	78	100.0	183	US-10-216-436-815	Sequence 815, Appl
14	78	100.0	287	US-09-791-537-102207	Sequence 102207, A
15	78	100.0	295	US-09-805-020-48	Sequence 48, Appl
16	78	100.0	300	PCT-US02-30474-1725	Sequence 1725, Ap
17	78	100.0	300	US-60-324-631-1730	Sequence 1730, Ap
18	78	100.0	317	US-09-791-537-125777	Sequence 125777, A
19	78	100.0	318	US-08-957-302-12	Sequence 12, Appl
20	78	100.0	318	US-09-522-046-2	Sequence 2, Appl
21	78	100.0	318	US-09-791-537-31856	Sequence 31856, A
22	78	100.0	318	US-09-791-537-61013	Sequence 61013, A
23	78	100.0	318	US-09-791-537-78158	Sequence 78158, A
24	78	100.0	318	US-09-791-537-79508	Sequence 79508, A
25	78	100.0	318	US-09-791-537-114574	Sequence 114574, A
26	78	100.0	318	US-09-976-594-938	Sequence 938, Ap
27	78	100.0	318	US-10-247-671-1374	Sequence 131, Ap
28	78	100.0	318	US-60-240-409-938	Sequence 938, Ap
29	78	100.0	318	US-60-323-784-131	Sequence 131, Ap
30	78	100.0	319	US-09-791-537-130991	Sequence 130991, A
31	78	100.0	319	US-09-791-537-67773	Sequence 67773, A
32	78	100.0	319	US-09-791-537-44105	Sequence 44105, A
33	78	100.0	319	US-09-791-537-114573	Sequence 114573, A
34	78	100.0	345	US-09-791-537-67718	Sequence 67718, A
35	78	100.0	345	US-09-802-540-12583	Sequence 12583, A
36	78	100.0	522	US-08-957-302-2	Sequence 145174, A
37	74	94.9	316	US-09-791-537-91767	Sequence 91767, A
38	74	94.9	317	US-09-791-537-8668	Sequence 8668, Ap
39	74	94.9	317	US-09-791-537-130991	Sequence 130991, A
40	64	82.1	219	US-09-791-537-67773	Sequence 67773, A
41	64	82.1	317	US-09-791-537-44105	Sequence 44105, A
42	62	79.5	318	US-09-791-537-114573	Sequence 114573, A
43	46	59.0	276	US-09-791-537-67718	Sequence 67718, A
44	46	59.0	577	US-09-802-540-12583	Sequence 12583, A
45	45	57.7	438	US-09-791-537-145174	Sequence 145174, A

#### ALIGNMENTS

RESULT 1  
US-09-440-597-2  
Sequence 2, Application US/09440597  
GENERAL INFORMATION:  
APPLICANT: KANAGUCHI, Haruma  
TITLE OF INVENTION: DRUG-IMMOBILIZED PARTICLES AND A PROCESS OF PURIFYING PROTEINS  
NUMBER OF SEQUENCE ADDRESSES: 13  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Mike, Bronstein, Roberts & Cushman, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/440,597  
FILING DATE: 15-Nov-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/795,927  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:

NAME: Corless, Peter F  
REGISTRATION NUMBER: 33,860  
REFERENCE/DOCKET NUMBER: 47115  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-440-597-2

Query Match 100.0%; Score 78; DB 18; Length 14;  
Best Local Similarity 100.0%; Pred. No. 2.7e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAGEGPALYEDPPD 14  
Db 1 AAGEGPALYEDPPD 14

RESULT 2  
US-10-001-426-2  
Sequence 2, Application US/10001426  
GENERAL INFORMATION:  
APPLICANT: HANDA, Hiroshi  
KAWAGUCHI, Haruma  
TITLE OF INVENTION: DRUG-IMMOBILIZED PARTICLES AND A  
PROCESS OF PURIFYING PROTEINS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/001,426  
FILING DATE: 02-Nov-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/440,597  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Corless, Peter F  
REGISTRATION NUMBER: 33,860  
REFERENCE/DOCKET NUMBER: 47115  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-001-426-2

Query Match 100.0%; Score 78; DB 24; Length 14;  
Best Local Similarity 100.0%; Pred. No. 2.7e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAGEGPALYEDPPD 14  
Db 1 AAGEGPALYEDPPD 14

RESULT 3  
US-10-003-014-2  
Sequence 2, Application US/10003014  
GENERAL INFORMATION:  
APPLICANT: HANDA, Hiroshi  
KAWAGUCHI, Haruma  
TITLE OF INVENTION: DRUG-IMMOBILIZED PARTICLES AND A  
PROCESS OF PURIFYING PROTEINS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/003,014  
FILING DATE: 02-Nov-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/440,597  
FILING DATE: 15-Nov-1999  
APPLICATION NUMBER: 08/795,927  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Corless, Peter F  
REGISTRATION NUMBER: 33,860  
REFERENCE/DOCKET NUMBER: 47115  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-003-014-2

Query Match 100.0%; Score 78; DB 24; Length 14;  
Best Local Similarity 100.0%; Pred. No. 2.7e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAGEGPALYEDPPD 14

Db 1 AAGEGPAlyEDPPD 14  
|||||

RESULT 4  
US-10-029-386-28638

; Sequence 28638, Application US/10029386  
; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G

; FILE REFERENCE: AEOMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029,386

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 34288

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 28638

; LENGTH: 69

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO CHR14.1

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.91

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2

; OTHER INFORMATION: SWISSPROT HIT: P27695, EVALU = 3.00e-24

US-10-029-386-28638

Query Match 100.0%; Score 78; DB 24; Length 69;

Best Local Similarity 100.0%; Pred. No. 0.00017;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGEGPAlyEDPPD 14  
|||||

Db 22 AAGEGPAlyEDPPD 35

RESULT 5

US-09-834-366-16628

; Sequence 16628, Application US/09834366  
; GENERAL INFORMATION:

; APPLICANT: Benjamin, Stephane

; APPLICANT: Tanaka, Hiroaki

; APPLICANT: Dumas Milne Edwards, Jean Baptiste

; APPLICANT: Jobert, Severin

; APPLICANT: Giordano, Jean-Yves

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: 81.US2.REG

; CURRENT APPLICATION NUMBER: US/09/834,366

; CURRENT FILING DATE: 2001-04-13

; PRIOR APPLICATION NUMBER: US 60/197,873

; PRIOR FILING DATE: 2000-04-18

; NUMBER OF SEQ ID NOS: 52153

; SOFTWARE: Patent.pm

; SEQ ID NO 16628

; LENGTH: 73

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: 59

; OTHER INFORMATION: Xaa = Pro,Thr

US-09-834-366-16628

Query Match 100.0%; Score 78; DB 22; Length 73;

Best Local Similarity 100.0%; Pred. No. 0.00018;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGEGPAlyEDPPD 14  
|||||

Db 37 AAGEGPAlyEDPPD 50

RESULT 6

US-60-197-873-16628

; Sequence 16628, Application US/60197873  
; GENERAL INFORMATION:

; APPLICANT: Benjamin, Stephane

; APPLICANT: Tanaka, Hiroaki

; APPLICANT: Dumas Milne Edwards, Jean Baptiste

; APPLICANT: Jobert, Severin

; APPLICANT: Giordano, Jean-Yves

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: 81.US1.PRO

; CURRENT APPLICATION NUMBER: US/60/197,873

; CURRENT FILING DATE: 2000-04-18

; NUMBER OF SEQ ID NOS: 52153

; SOFTWARE: Patent.pm

; SEQ ID NO 16628

; LENGTH: 73

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: 59

; OTHER INFORMATION: Xaa = Pro,Thr

US-60-197-873-16628

Query Match 100.0%; Score 78; DB 27; Length 73;

Best Local Similarity 100.0%; Pred. No. 0.00018;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGEGPAlyEDPPD 14  
|||||

Db 37 AAGEGPAlyEDPPD 50

RESULT 7

US-09-791-537-41381

; Sequence 41381, Application US/09791537  
; GENERAL INFORMATION:

; APPLICANT: Bionomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 41381

; LENGTH: 146

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-791-537-41381

Query Match 100.0%; Score 78; DB 21; Length 146;

Best Local Similarity 100.0%; Pred. No. 0.00041;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGEGPAlyEDPPD 14  
|||||

Db 37 AAGEGPAlyEDPPD 50

RESULT 8

PCT-US01-18569-4083

; Sequence 4083, Application PC/TUS0118569  
; GENERAL INFORMATION:

US-09-834-366-16628

Query Match 100.0%; Score 78; DB 22; Length 73;

Best Local Similarity 100.0%; Pred. No. 0.00018;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA133PCT
; CURRENT APPLICATION NUMBER: PCT/US01/18569
; CURRENT FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4083
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (61)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-18569-4083
```

```
Query Match          100.0%; Score 78; DB 1; Length 170;
Best Local Similarity 100.0%; Pred. No. 0.00049;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AAGEGPALYEDPPD 14
Db 68 AAGEGPALYEDPPD 81
```

```
RESULT 9
US-10-264-049-4083
; Sequence 4083, Application US/10264049
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PCT
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 4083
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (61)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-4083
```

```
Query Match          100.0%; Score 78; DB 26; Length 170;
Best Local Similarity 100.0%; Pred. No. 0.00049;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AAGEGPALYEDPPD 14
Db 68 AAGEGPALYEDPPD 81
```

```
RESULT 10
US-09-760-482-73
; Sequence 73, Application US/09760482
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT260
; CURRENT APPLICATION NUMBER: US/09/760,482
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 215
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (175)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; LOCATION: (182)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-760-482-73
```

```
Query Match          100.0%; Score 78; DB 21; Length 183;
Best Local Similarity 100.0%; Pred. No. 0.00053;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AAGEGPALYEDPPD 14
Db 64 AAGEGPALYEDPPD 77
```

```
RESULT 11
US-09-760-485-815
; Sequence 815, Application US/09760485
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ217
; CURRENT APPLICATION NUMBER: US/09/760,485
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1477
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 815
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (175)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (182)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-760-485-815
```

```
Query Match          100.0%; Score 78; DB 21; Length 183;
Best Local Similarity 100.0%; Pred. No. 0.00053;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AAGEGPALYEDPPD 14
Db 64 AAGEGPALYEDPPD 77
```

```
RESULT 12
US-10-153-088-73
; Sequence 73, Application US/10153088
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT260CIN
; CURRENT APPLICATION NUMBER: US/10/153,088
; CURRENT FILING DATE: 2002-05-23
; Prior application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (175)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (182)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-153-086-73
```

```
Query Match          100.0%; Score 78; DB 25; Length 183;
Best Local Similarity 100.0%; Pred. No. 0.00053;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AGECPALYEDPPD 14
DB 64 AGECPALYEDPPD 77
```

## RESULT 13

```
US-10-216-436-815
; Sequence 815, Application US/10216436
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0217C1N
; CURRENT APPLICATION NUMBER: US/10/216,436
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/760,485
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1477
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 815
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (175)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (182)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-216-436-815
```

```
Query Match          100.0%; Score 78; DB 26; Length 183;
Best Local Similarity 100.0%; Pred. No. 0.00053;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AGECPALYEDPPD 14
DB 64 AGECPALYEDPPD 77
```

```
RESULT 14
US-09-791-537-102207
; Sequence 102207, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 102207
; LENGTH: 287
; TYPE: PRT
; ORGANISM: pdb 1BIX
US-09-791-537-102207
```

```
Query Match          100.0%; Score 78; DB 21; Length 287;
Best Local Similarity 100.0%; Pred. No. 0.00089;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AGECPALYEDPPD 14
DB 6 AGECPALYEDPPD 19
```

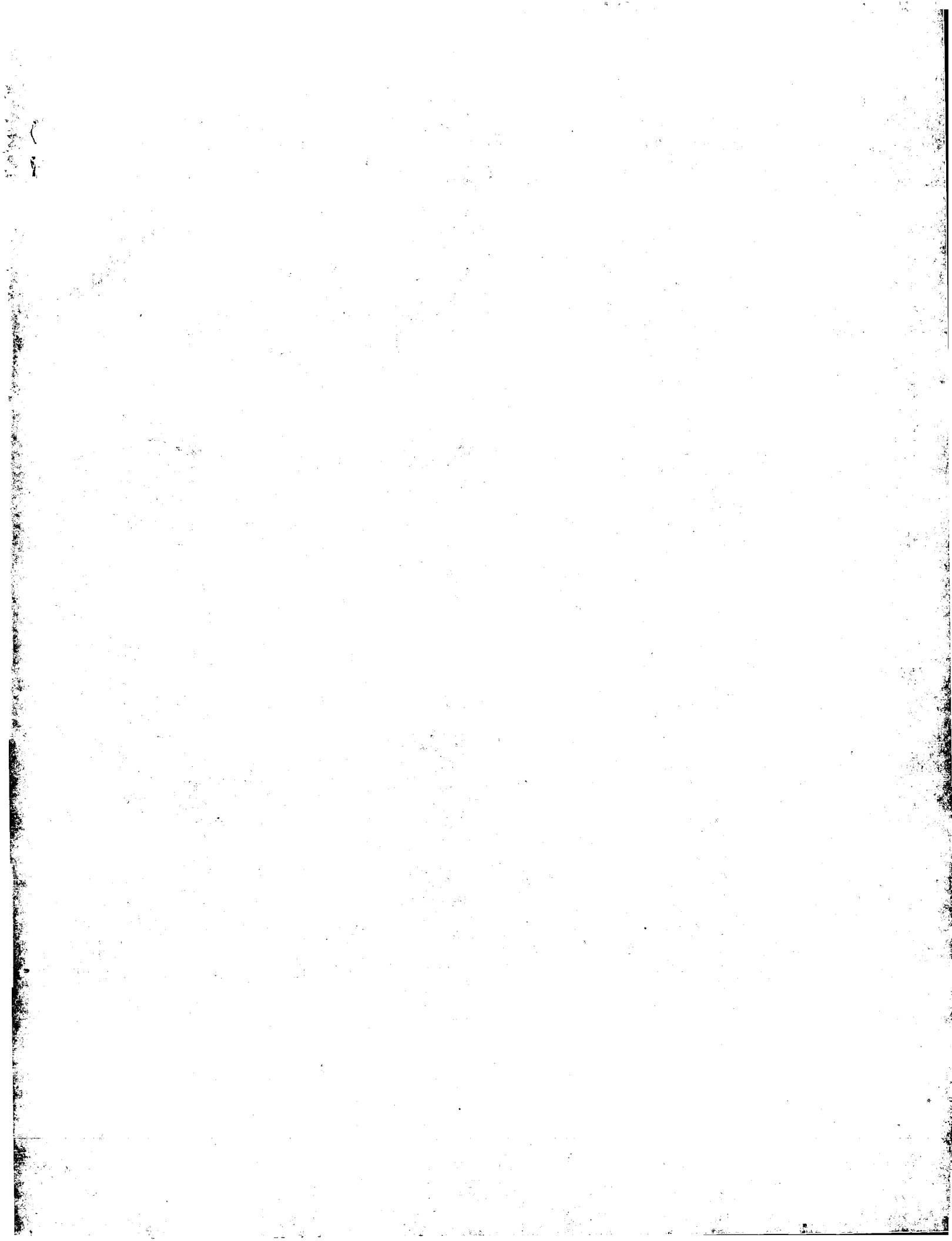
## RESULT 15

```
US-09-805-020-48
; Sequence 48, Application US/09805020
; GENERAL INFORMATION:
; APPLICANT: LEVINE, Zurilt
; TITLE OF INVENTION: SPLICE VARIANTS OF ONCOGENES
; FILE REFERENCE: 2786-0168P
; CURRENT APPLICATION NUMBER: US/09/805,020
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 48
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-805-020-48
```

```
Query Match          100.0%; Score 78; DB 22; Length 295;
Best Local Similarity 100.0%; Pred. No. 0.00092;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AGECPALYEDPPD 14
DB 37 AGECPALYEDPPD 50
```

Search completed: January 30, 2003, 14:41:41  
Job time : 141 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2003, 14:38:04 : Search time 18 seconds  
(without alignments)  
61.723 Million cell updates/sec

Title: US-10-001-426-2  
Perfect score: 78  
Sequence: 1 AAGEGPALEDPPD 14

Scoring table:  
BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 398253 seqs, 79358460 residues

Total number of hits satisfying chosen parameters: 398253

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New : \*  
1: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	282	5	US-09-724-676-91737 Sequence 91737, A
2	78	100.0	282	5	US-09-724-676-91739 Sequence 91739, A
3	78	100.0	282	5	US-09-724-676-91737 Sequence 91737, A
4	78	100.0	282	5	US-09-724-676-91739 Sequence 91739, A
5	78	100.0	293	5	US-09-724-676-91745 Sequence 91745, A
6	78	100.0	293	5	US-09-724-676-91746 Sequence 91746, A
7	78	100.0	293	5	US-09-724-676-91758 Sequence 91758, A
8	78	100.0	293	5	US-09-724-676-91759 Sequence 91759, A
9	78	100.0	293	5	US-09-724-676-91769 Sequence 91769, A
10	78	100.0	293	5	US-09-724-676-91770 Sequence 91770, A
11	78	100.0	293	5	US-09-724-676-91745 Sequence 91745, A
12	78	100.0	293	5	US-09-724-676-91758 Sequence 91758, A
13	78	100.0	293	5	US-09-724-676-91759 Sequence 91759, A
14	78	100.0	293	5	US-09-724-676-91769 Sequence 91769, A
15	78	100.0	293	5	US-09-724-676-91770 Sequence 91770, A
16	78	100.0	293	5	US-09-724-676-91741 Sequence 91741, A
17	78	100.0	293	5	US-09-724-676-91742 Sequence 91742, A
18	78	100.0	293	5	US-09-724-676-91749 Sequence 91749, A
19	78	100.0	293	5	US-09-724-676-91755 Sequence 91755, A
20	78	100.0	293	5	US-09-724-676-91760 Sequence 91760, A
21	78	100.0	293	5	US-09-724-676-91741 Sequence 91741, A
22	78	100.0	293	5	US-09-724-676-91742 Sequence 91742, A
23	78	100.0	293	5	US-09-724-676-91749 Sequence 91749, A
24	78	100.0	293	5	US-09-724-676-91754 Sequence 91754, A
25	78	100.0	293	5	US-09-724-676-91754 Sequence 91754, A
26	78	100.0	293	5	US-09-724-676-91754 Sequence 91754, A

27	78	100.0	295	5	US-09-724-676A-91755 Sequence 91755, A
28	78	100.0	295	5	US-09-724-676A-91760 Sequence 91760, A
29	78	100.0	298	5	US-09-724-676-91733 Sequence 91733, A
30	78	100.0	298	5	US-09-724-676-91734 Sequence 91734, A
31	78	100.0	298	5	US-09-724-676A-91733 Sequence 91733, A
32	78	100.0	298	5	US-09-724-676A-91734 Sequence 91734, A
33	78	100.0	305	5	US-09-724-676-91735 Sequence 91735, A
34	78	100.0	305	5	US-09-724-676-91736 Sequence 91736, A
35	78	100.0	305	5	US-09-724-676A-91735 Sequence 91735, A
36	78	100.0	316	5	US-09-724-676A-91736 Sequence 91736, A
37	78	100.0	316	5	US-09-724-676-91743 Sequence 91743, A
38	78	100.0	316	5	US-09-724-676-91744 Sequence 91744, A
39	78	100.0	316	5	US-09-724-676-91756 Sequence 91756, A
40	78	100.0	316	5	US-09-724-676-91757 Sequence 91757, A
41	78	100.0	316	5	US-09-724-676-91767 Sequence 91767, A
42	78	100.0	316	5	US-09-724-676-91768 Sequence 91768, A
43	78	100.0	316	5	US-09-724-676A-91743 Sequence 91743, A
44	78	100.0	316	5	US-09-724-676A-91744 Sequence 91744, A
45	78	100.0	316	5	US-09-724-676A-91756 Sequence 91756, A

## ALIGNMENTS

```

RESULT 1
US-09-724-676-91737
; Sequence 91737, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 91737
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-91737

Query Match      100.0%; Score 78; DB 5; Length 282;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAGEGPALEDPPD 14
      |||
DB      24 AAGEGPALEDPPD 37
      |||

RESULT 2
US-09-724-676-91739
; Sequence 91739, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 91739
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-91739

Query Match      100.0%; Score 78; DB 5; Length 282;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAGEGPALEDPPD 14
      |||
  
```

Db 24 AAGEGPALYEDPPD 37

RESULT 3

US-09-724-676A-91737  
; Sequence 91737, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 91737  
; LENGTH: 282  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-91737

Query Match

Best Local Similarity 100.0%; Score 78; DB 5; Length 282;  
Pred. No. 1.6e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGEGPALYEDPPD 14

Db 24 AAGEGPALYEDPPD 37

RESULT 4

US-09-724-676A-91739  
; Sequence 91739, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 91739  
; LENGTH: 282  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-91739

Query Match

Best Local Similarity 100.0%; Score 78; DB 5; Length 282;  
Pred. No. 1.6e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGEGPALYEDPPD 14

Db 24 AAGEGPALYEDPPD 37

RESULT 5

US-09-724-676-91745  
; Sequence 91745, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 91745  
; LENGTH: 293  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-91745

Query Match

Best Local Similarity 100.0%; Score 78; DB 5; Length 293;

Best Local Similarity 100.0%; Pred. No. 1.7e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGEGPALYEDPPD 14

Db 35 AAGEGPALYEDPPD 48

RESULT 6

US-09-724-676-91746  
; Sequence 91746, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 91746  
; LENGTH: 293  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-91746

Query Match

Best Local Similarity 100.0%; Score 78; DB 5; Length 293;  
Pred. No. 1.7e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGEGPALYEDPPD 14

Db 35 AAGEGPALYEDPPD 48

RESULT 7

US-09-724-676-91758  
; Sequence 91758, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 91758  
; LENGTH: 293  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-91758

Query Match

Best Local Similarity 100.0%; Score 78; DB 5; Length 293;  
Pred. No. 1.7e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGEGPALYEDPPD 14

Db 35 AAGEGPALYEDPPD 48

RESULT 8

US-09-724-676-91759  
; Sequence 91759, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 91759  
; LENGTH: 293

Query Match

Best Local Similarity 100.0%; Score 78; DB 5; Length 293;

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-91759

```

```

Query Match
Best Local Similarity 100.0%; Score 78; DB 5; Length 293;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 AAGEGPALEYDDPD 14
Db 35 AAGEGPALEYDDPD 48

```

```

RESULT 9
US-09-724-676-91769
; Sequence 91769, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 91769
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-91769

```

```

Query Match
Best Local Similarity 100.0%; Score 78; DB 5; Length 293;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 AAGEGPALEYDDPD 14
Db 35 AAGEGPALEYDDPD 48

```

```

RESULT 10
US-09-724-676-91770
; Sequence 91770, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 91770
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-91770

```

```

Query Match
Best Local Similarity 100.0%; Score 78; DB 5; Length 293;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 AAGEGPALEYDDPD 14
Db 35 AAGEGPALEYDDPD 48

```

```

RESULT 11
US-09-724-676A-91745
; Sequence 91745, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A

```

```

; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 91745
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-91745

```

```

Query Match
Best Local Similarity 100.0%; Score 78; DB 5; Length 293;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 AAGEGPALEYDDPD 14
Db 35 AAGEGPALEYDDPD 48

```

```

RESULT 12
US-09-724-676A-91746
; Sequence 91746, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 91746
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-91746

```

```

Query Match
Best Local Similarity 100.0%; Score 78; DB 5; Length 293;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 AAGEGPALEYDDPD 14
Db 35 AAGEGPALEYDDPD 48

```

```

RESULT 13
US-09-724-676A-91758
; Sequence 91758, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 91758
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-91758

```

```

Query Match
Best Local Similarity 100.0%; Score 78; DB 5; Length 293;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 AAGEGPALEYDDPD 14
Db 35 AAGEGPALEYDDPD 48

```

```

RESULT 14
US-09-724-676A-91759
; Sequence 91759, Application US/09724676A

```

```
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; NUMBER OF SEQ ID NOS: 2000-11-28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 91759
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-91759

Query Match          100.0%; Score 78; DB 5; Length 293;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGEGPALYEDPPD 14
   |||||||
Db 35 AAGEGPALYEDPPD 48

RESULT 15
US-09-724-676A-91769
; Sequence 91769, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; NUMBER OF SEQ ID NOS: 2000-11-28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 91769
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-91769

Query Match          100.0%; Score 78; DB 5; Length 293;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Search completed: January 30, 2003, 14:42:05  
Job time : 19 secs

Thu Jan 30 17:27:57 2003

us-10-001-426-2.rag

NEFF  
10/001, 426

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2003, 14:23:48 ; Search time 35 Seconds  
(without alignments)  
53.300 Million cell updates/sec

Title: US-10-001-426-2

Perfect score: 78  
Sequence: 1 AAGCPALYEDPPD 14

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

```
1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	14	AAW21748	Fragment #2 of E33
2	78	100.0	170	ABP42951	Human ovarian anti
3	78	100.0	318	AAW21746	E3330-binding prot
4	78	100.0	318	AAW48894	Homo sapiens AP en
5	78	100.0	318	AAW52864	Apurinic/apyrimidi
6	78	100.0	345	AAW56610	Human prostate can
7	78	100.0	324	AAW48893	Homo sapiens MGMT-
8	45	57.7	364	AAW53366	Protonibacterium
9	44	56.4	68	AAW54099	Human pancreatic c
10	43	55.1	241	AAW47601	Protonibacterium

11	43	55.1	999	21	AAG53685
12	43	55.1	1008	21	AAG53684
13	43	55.1	1065	21	AAG53683
14	42	53.8	115	19	AAV21361
15	42	53.8	383	23	ABB97554
16	42	53.8	384	22	AAG67464
17	42	53.8	583	22	ABB71010
18	42	53.8	1017	22	ABB95217
19	42	53.8	1146	20	AAW96322
20	42	53.8	1153	15	AAW57664
21	42	53.8	1153	15	AAW63206
22	42	53.8	1153	17	AAW88464
23	42	53.8	1153	19	AAW36113
24	42	53.8	1153	22	AAW64497
25	42	53.8	1153	22	AAW64498
26	42	53.8	1153	22	AAW31724
27	42	53.8	1153	22	AAW66724
28	42	53.8	1153	22	AAW67463
29	42	53.8	1182	23	AAW19798
30	42	53.8	1207	23	AAW19797
31	41	52.6	107	22	AAW51428
32	41	52.6	314	22	AAW94555
33	41	52.6	1144	16	AAW77360
34	41	52.6	1144	19	AAW51246
35	41	52.6	1144	22	AAW64500
36	41	52.6	1147	17	AAW02571
37	41	52.6	1189	23	AAW19796
38	41	52.6	1319	22	AAW22751
39	40	51.3	65	22	AAW31478
40	40	51.3	100	22	AAW31479
41	40	51.3	134	21	AAW41036
42	40	51.3	134	23	AAW11447
43	40	51.3	491	22	AAW39422
44	40	51.3	825	15	AAW60811
45	40	51.3	904	22	ABG20852

#### ALIGNMENTS

```
RESULT 1
AAW21748
ID AAW21748 standard; peptide: 14 AA.
AC AAW21748;
DT 02-MAR-1998 (first entry)
DE Fragment #2 of E3330-binding protein, Ref-1.
KW E3330-binding protein; Ref-1; microsphere; drug binding factor recovery;
KW styrene-glycidyl methacrylate polymer; protein isolation;
KW protein purification; receptor identification.
XX
XX
XX Synthetic.
OS
XX
XX EP787988-A2.
PN
XX
XX 06-AUG-1997.
PD
XX
XX 05-FEB-1997; 97EP-0101821.
PF
XX
XX 17-SEP-1996; 96CP-0266711.
PR
XX
XX 05-FEB-1996; 96CP-0018827.
PA
XX
XX (HAND/) HANDA H.
PA (KAWA/) KAWAGUCHI H.
PI
XX
XX Handa H, Kawaguchi H;
PI
XX
XX WPI; 1997-387634/36.
DR
XX
XX Isolation and identification of receptors to specific compounds -
PT
```

PT using microspheres prepared by coupling the compound, via a spacer, to a styrene-glycidyl methacrylate polymer.

PS Example 9; Page 10; 29pp; English

CC AAM2174.W121749 represent fragments of the E3330-binding protein, Ref-1  
CC (see AAM21746). E3330 is  
CC 3-(1-(5-(2,3-dimethoxy-6-methyl-1,4-benzoquinonyl))-2-nonyl-2-propionic  
CC acid. Ref-1 was isolated using the method of the invention. The method of  
CC the invention is for isolating a substance (S1) that can adhere to a  
CC substance (S2) possessing physiological activity from a mixture  
CC containing S1. The method comprises contacting the mixture with a  
CC microsphere prepared by coupling S2 to a styrene-glycidyl methacrylate  
CC polymer through a spacer. The process is useful, e.g., for recovery of  
CC binding factors (such as proteins) from cell extracts. The process  
CC allows easy isolation, purification and identification of receptors to  
CC specific compounds.

50 Sequence 14 AA;

Query Match	100.0%	Score 78;	DB 18;	Length 14;
Best Local Similarity	100.0%;	Pred. No. 1.1e-05;		
Matches 14; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1	AAGEGPALYEDPPD	14
Db	1	AAGEGPALYEDPPD	14

Db 1 AAGGPGALYEDPPD 14

RESULT 2  
ABP42951  
ID ABP42951 standard; Protein; 170 AA

... AC ABP42951;

DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HPDVI25, SEQ ID NO:4083.

KM Human, ovarian antigen; ovary; ovarian; breast, cancer; tumour;  
 KM ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KM infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KM PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KM inflammatory condition; immune disorder; blood disorder;  
 KM cardiovascular disorder; respiratory system disorder; neurological disorder;  
 KM gastrointestinal disorder; urinary system disorder; drug screening;  
 KM gene therapy; chromosome mapping; forensic analysis;  
 KM antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KM antiinflammatory; gynaecological; reproductive.

OS Homo sapiens.

PN WQ200200677-A1

03 - JAN - 2002 .  
PD

07-JUN-2001; 2001WO-US18569.

07-JUN-2000; 2000US-209467P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;

DR WPI; 2002-147878/19.

DR N-PSDB; ABQ56028.

PT Isolated nucleic acid molecules encoding novel ovarian polypeptides  
PT useful in the prevention, treatment and diagnosis of cancer (e.g.  
PT ovarian cancer), immune disorders, cardiovascular disorders and  
PT neurological diseases -

PS Claim 11; SEQ ID No 4083; 2922pp; English

xx The invention relates to 2175 novel human ovarian antigens (ABP41054-  
cc ABP43228) and to cDNAs encoding them (AB054131-AB056305), and also  
cc encompasses polypeptides 90% identical and polynucleotides 95% identical  
cc to the sequences of the invention. The invention additionally relates to  
cc recombinant vectors and host cells comprising human ovarian antigen  
cc polynucleotides, antibodies against human ovarian antigens, and the use  
cc of ovarian antigen polynucleotides and polypeptides in diagnosing,  
cc treating, prognosing or preventing various ovarian and/or breast-related  
cc disorders. Such conditions include ovarian cancer and breast cancer, and  
cc metastatic tumours of ovarian or breast origin, reproductive system  
cc disorders (e.g., infertility, disorders of pregnancy, anovulation,  
cc polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
cc disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
cc shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
cc vaginitis), immune disorders (e.g., congenital and acquired  
cc immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
cc blood-related disorders (e.g., anaemia), cardiovascular disorders,  
cc respiratory disorders, neurological disorders, gastrointestinal disorders  
cc and urinary system disorders. Ovarian antigen polypeptides and  
cc polynucleotides may also be used in screening for compounds which  
cc modulate ovarian antigen expression or activity. The polynucleotides may  
cc further be used for gene therapy, chromosome mapping, in the  
cc identification of individuals and in forensic analysis, and the  
cc polypeptides may be used as food additives or to prepare antibodies  
cc useful in disease diagnosis, drug targeting and phenotyping. The present  
cc sequence represents a human ovarian antigen of the invention.  
cc Note: The sequence data for this patent did not form part of the printed  
cc specification, but was obtained in electronic format directly from WIPO  
cc at ftp.wipo.int/pub/published\_pt\_sequences.

Sequence 170 AA;

Query Match	100.0%;	Score 78;	DB 23;	Length 170;
Best Local Similarity	100.0%;	Pred. No. 0.00014;		
Matches 14; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      1 AAGEGPALYEDPPD 14
          |||||
Db      68 AAGEGPALYEDPPD 81
```

Db 68 AAGEGPALYEDPPD 81

RESULT 3  
AAW21746  
ID AAW21746 standard; protein; 318 AA

AC AAW21746;

DT 02-MAR-1998 (first entry)

DE E3330-binding protein, Ref-1.

\*\*\*  
KM E3330-binding protein; Ref-1; microsphere; drug binding factor recovery;  
KM styrene-glycidyl methacrylate polymer; protein isolation;  
KM protein purification; receptor identification.

OS Synthetic

PN EP787988-A2.

06-AUG-1997 PD

05-FEB-1997; 97EP-0101821.

PR 17-SEP-1996; 96JP-0266711.

PR 05-FEB-1996; 96JP-0018827

PA (HAND/) HANDA H.  
PA (KAWA/) KAWAGUCHI H.

PI Handa H, Kawaguchi H;

DR WPI; 1997-387634/36.

XX Isolation and identification of receptors to specific compounds -  
 PT using microspheres prepared by coupling the compound, via a spacer,  
 PT to a styrene-glycidyl methacrylate polymer.  
 XX  
 XX Claim 5; Page 11; 29pp; English.  
 CC This sequence represents the E3330-binding protein, Ref-1. E3330 is  
 CC 3-[(5-(2,3-dimethoxy-6-methyl-1,4-benzoquinonyl)-2-nonyl-2-propionyl  
 CC acid. This sequence was isolated using the method of the invention. The  
 CC method of the invention is for isolating a substance (S1) that can adhere  
 CC to a substance (S2) possessing physiological activity from a mixture  
 CC containing S1. The method comprises contacting the mixture with a  
 CC microsphere prepared by coupling S2 to a styrene-glycidyl methacrylate  
 CC polymer through a spacer. The process is useful, e.g., for recovery of  
 CC drug binding factors (such as proteins) from cell extracts. The process  
 CC allows easy isolation, purification and identification of receptors to  
 CC specific compounds.  
 XX  
 SO Sequence 318 AA;  
 Query Match 100.0%; Score 78; DB 18; Length 318;  
 Best Local Similarity 100.0%; Pred. No. 0.00026;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 AAGEGPAlyEDPPD 14  
 DB 37 AAGEGPAlyEDPPD 50  
 RESULT 4  
 AAM48894  
 ID AAM48894 standard; Protein: 318 AA.  
 XX  
 AC AAM48894;  
 XX  
 DT 13-OCT-1998 (first entry)  
 XX  
 DE Homo sapiens AP endonuclease.  
 XX  
 KM MGMT; APE; DNA repair protein; fusion protein; AP endonuclease;  
 KM 0-6-methylguanine-DNA methyltransferase.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9817684-A2.  
 XX  
 PD 30-APR-1998.  
 XX  
 PE 24-OCT-1997; 97WO-US19629.  
 XX  
 PR 25-OCT-1996; 96US-0029308.  
 XX  
 PA (ADRE-) ADVANCED RES & TECHNOLOGY INST.  
 XX  
 PI Kelley M, Williams D;  
 XX  
 DR WPI: 1998-271785/24.  
 DR N-PSDB: AAV32543.  
 XX  
 PT New fusion proteins comprising 2 DNA repair proteins - used for  
 PT protecting cells against DNA damage, e.g. by chemotherapeutic  
 PT agents, or for treating DNA-repair-defective human diseases  
 XX  
 PS Disclosure; p115-116; 150pp; English.  
 XX  
 CC The sequence is that encoding the APE gene which was used in the  
 CC construction of a fusion protein of the DNA repair  
 CC proteins 0-6-methylguanine-DNA methyltransferase (MGMT) and AP  
 CC endonuclease (APE). It can be used for protecting cells  
 CC (especially bone marrow cells) against DNA damaging agents such as  
 CC chemotherapeutic agents used in tumour therapy and can be used in  
 CC conjunction with chemotherapy to facilitate the use of increased doses

CC of chemotherapy. They can provide added protection to cells by  
 CC providing RPS which recognise different types of DNA lesions. In  
 CC addition, the RPS can be used in the treatment of DNA-repair-defective  
 CC human diseases, e.g. xeroderma pigmentosum, Ataxia telangiectasia,  
 CC Cockayne's syndrome, Bloom's syndrome, Fanconi's anemia, Parkinson's  
 CC disease and other neurodegenerative diseases including amyotrophic  
 CC lateral sclerosis, Alzheimer's disease and other age related disorders.  
 XX  
 SO Sequence 318 AA;  
 Query Match 100.0%; Score 78; DB 19; Length 318;  
 Best Local Similarity 100.0%; Pred. No. 0.00026;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 AAGEGPAlyEDPPD 14  
 DB 37 AAGEGPAlyEDPPD 50  
 RESULT 5  
 AAM52864  
 ID AAM52864 standard; Protein: 318 AA.  
 XX  
 AC AAM52864;  
 XX  
 DT 08-JUL-1998 (first entry)  
 XX  
 DE Apurinic/apurymidinic endonuclease amino acid sequence.  
 XX  
 KM Apurinic/apurymidinic endonuclease; APE; identification; premalignancy;  
 KM APE level; diagnosis; therapy; tumour; apoptosis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9747971-A1.  
 XX  
 PD 18-DEC-1997.  
 XX  
 PE 11-JUN-1997; 97WO-US10078.  
 XX  
 PR 11-JUN-1996; 96US-0019602.  
 PR 11-JUN-1996; 96US-0019561.  
 XX  
 PA (ADRE-) ADVANCED RES & TECHNOLOGY INST.  
 XX  
 PI Duguid JR, Ebler JN, Kelley MR;  
 XX  
 DR WPI: 1998-168751/15.  
 DR N-PSDB: AAV21204.  
 XX  
 PT Use of apurinic/apurymidinic endonuclease - for developing products  
 PT for the diagnosis and therapy of tumours, or for detecting, inducing  
 PT or inhibiting apoptosis in cells  
 XX  
 PS Disclosure; Pages 131-132; 166pp; English.  
 XX  
 CC The present sequence represents an apurinic/apurymidinic endonuclease  
 CC (APE). Levels of APE in dysplasia, carcinoma in situ and squamous cell  
 CC carcinomas of the cervix and prostate are substantially elevated,  
 CC relative to controls. A method for identifying a premalignant or  
 CC malignant condition in a human subject comprises determining the  
 CC level of APE in cells from the subject, where an elevated level of  
 CC APE, as compared to the APE level in corresponding normal cells,  
 CC indicates a premalignant or malignant condition. APE levels can be  
 CC used in the diagnosis and therapy of tumours. They can also be used  
 CC for detecting, inducing or inhibiting apoptosis in cells, e.g. for  
 CC treating T-cells infected with HIV or detecting apoptosis in cells  
 CC that have been subjected to chemo-, radio- or gene therapy.  
 XX  
 SO Sequence 318 AA;  
 Query Match 100.0%; Score 78; DB 19; Length 318;  
 Best Local Similarity 100.0%; Pred. No. 0.00026;

Matches 14: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGEGPALEYDDPD 14  
 DB 37 AAGEGPALEYDDPD 50

## RESULT 6

AAB56610  
 ID AAB56610 standard; Protein; 345 AA.

AC AAB56610;

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen protein sequence SPO ID NO:1188.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
 KW neuroprotective; cytosolic; cardiostatic; immunomodulatory; muscular;  
 KW vulnary; gastrointestinal; nephrotoxic; anti-infective; gynaecological;  
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;  
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
 XX wound; infectious disease.

OS Homo sapiens.

PN W020005174-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05988.

PR 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.

PI Rosen CA, Ruben SM;

DR WPI; 2000-587513/55.

DR N-PSDB; AAF15813.

XX Prostate cancer associated gene sequences, referred to as prostate  
 PT cancer antigens, useful for treatment, prevention, and diagnosis of  
 PT disorders such as prostate cancer -

PS Claim 11; Page 1600; 2338pp; English.

XX AAF1566 to AAF16505 encode the human prostate cancer associated  
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.  
 CC The prostate cancer antigens can have neuroprotective, cytosolic,  
 CC cardiostatic, immunomodulatory, muscular, vulnary, gastrointestinal,  
 CC nephrotoxic, anti-infective, gynaecological and antibacterial activities,  
 CC and can be used in gene therapy. The prostate cancer antigen  
 CC polynucleotides may be used for detection of prostate cancer, chromosome  
 CC identification, as chromosome markers, and for numerous other diagnostic  
 CC or research purposes. The prostate cancer antigens may be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to  
 CC AAB57303 represent sequences used in the exemplification of the present  
 CC invention.

XX Sequence 345 AA;

Query Match 100.0%; Score 78; DB 21; Length 345;

Best Local Similarity 100.0%; Pred. No. 0.00028;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGEGPALEYDDPD 14  
 DB 64 AAGEGPALEYDDPD 77

RESULT 7  
 ID AAM48893 standard; Protein; 524 AA.

AC AAM48893;

DT 13-OCT-1998 (first entry)

DE Homo sapiens MGMT-APE fusion protein.

XX MGMT; APE; DNA repair protein; fusion protein; AP endonuclease;  
 KW 0-6-methylguanine-DNA methyltransferase.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Protein 1..207 /note="MGMT"

FT Protein 208..524 /note="APE"

PN W09817684-A2.

PD 30-APR-1998.

PF 24-OCT-1997; 97WO-US19629.

PR 25-OCT-1996; 96US-0029308.

XX (ADRE-) ADVANCED RES & TECHNOLOGY INST.

PI Kelley M, Williams D;

DR WPI: 1998-271785/24.

DR N-PSDB; AAV32540.

XX New fusion proteins comprising 2 DNA repair proteins - used for  
 PT protecting cells against DNA damage, e.g. by chemotherapeutic  
 PT agents, or for treating DNA-repair-defective human diseases

PS Claim 11; p108-110; 150pp; English.

XX The sequence is that of fusion protein of the DNA repair  
 CC proteins 0-6-methylguanine-DNA methyltransferase (MGMT) and AP  
 CC endonuclease (APE). It can be used for protecting cells  
 CC (especially bone marrow cells) against DNA damaging agents such as  
 CC chemotherapeutic agents used in tumour therapy and can be used in  
 CC conjunction with chemotherapy to facilitate the use of increased doses  
 CC of chemotherapy. They can provide added protection to cells by  
 CC providing RPs which recognise different types of DNA lesions. In  
 CC addition, the FPs can be used in the treatment of DNA-repair-defective  
 CC human diseases, e.g. Xeroderma Pigmentosum, Ataxia Telangiectasia,  
 CC Cockayne's syndrome, Bloom's syndrome, Fanconi's anaemia, Parkinson's  
 CC disease and other neurodegenerative diseases including amyotrophic  
 CC lateral sclerosis, Alzheimer's disease and other age related disorders.

XX Sequence 524 AA;

Query Match 100.0%; Score 78; DB 19; Length 524;

Best Local Similarity 100.0%; Pred. No. 0.00042;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGEGPALEYDDPD 14  
 DB 243 AAGEGPALEYDDPD 256

## RESULT 8

AAU54366  
 ID AAU54366 standard; Protein; 364 AA.

AC AAU54366;



```

XX 27-FEB-2002 (first entry)
XX
XX Proionibacterium acnes immunogenic protein #15262.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
XX Proionibacterium acnes.
XX
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US12865.
XX
XX 21-APR-2000; 2000US-199047P.
XX 02-JUN-2000; 2000US-208841P.
XX 07-JUL-2000; 2000US-216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX WPI: 2001-616774/71.
XX N-PSDB: AAS59564.
XX
XX Proionibacterium acnes polypeptides and nucleic acids useful for
XX vaccinating against and diagnosing infections, especially useful for
XX treating acne vulgaris -
XX
XX Example 1: SEQ ID No 15561: 1069pp; English.
XX
XX Sequences AAU9105-AAU68017 represent Proionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
XX P. acnes is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of P. acnes in a patient comprises contacting a
XX sample with a binding agent that binds to the proteins of the invention
XX and determining the amount of bound protein in the sample. The
XX polypeptides may be used as antigens in the production of antibodies
XX specific for P. acnes proteins. These antibodies can be used to
XX downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme linked immunosorbent assay (ELISA).
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_ptc_sequences.
XX
XX Sequence 364 AA:
XX
XX Query Match 57.7%; Score 45; DB 22; Length 364;
XX Best Local Similarity 66.7%; Pred. No. 45;
XX Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX Oy 3 GEGPALYEDPPD 14
XX | | | | |
XX | | | | |
XX Db 200 GESPVTGEDPPD 211
XX
XX RESULT 9
XX ID AAB54099
XX AAB54099 standard; Protein: 68 AA.
XX
XX AAB54099;

```

```

XX 09-MAR-2001 (first entry)
XX
XX Human pancreatic cancer antigen protein sequence SEQ ID NO:551.
XX
XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
XX detection; diagnosis; identification; cytostatic; neuroprotective;
XX neurotropic; immunomodulatory; relaxant; contraceptive; gynaecological;
XX antiinflammatory; cardiant; gene therapy; chromosome mapping;
XX linkage analysis; tissue identification; tissue typing; forensic;
XX neural; immune system; muscular; reproductive; gastrointestinal;
XX pulmonary; cardiovascular; renal; proliferative.
XX
XX Homo sapiens.
XX
XX WO200055320-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05989.
XX
XX 12-MAR-1999; 99US-0124270.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX WPI: 2000-579444/54.
XX N-PSDB: AAC98864.
XX
XX New nucleic acid that is a pancreatic cancer antigen for preventing,
XX treating, or ameliorating a medical condition, particular pancreatic
XX cancer, or for use in assays for diagnosing a pathological condition -
XX
XX Claim 11: Page 989; 1379pp; English.
XX
XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
XX proteins, called pancreatic cancer antigens, given in AAB54008 to
XX AAB54466. The human pancreatic cancer antigens have cytostatic,
XX neuroprotective, neurotropic, immunomodulatory, relaxant, contraceptive,
XX gynaecological, cardiant and antiinflammatory activities, and can be used
XX in gene therapy. The polynucleotide and proteins can be used for
XX preventing, treating, or ameliorating a medical condition or in assays
XX for diagnosing a pathological condition or a susceptibility to one in a
XX subject. Binding partners to the proteins and the activity of the
XX proteins can be identified. The pancreatic cancer antigens may be used to
XX detect, treat or prevent pancreatic disorders, especially cancer.
XX Agonists and antagonists to the antigens can be screened for. The
XX pancreatic cancer antigen polynucleotides can be used to design nucleic
XX acid hybridisation probes that can be used in chromosome mapping, linkage
XX analysis, tissue identification and/or typing and a variety of forensic
XX and diagnostic methods. The proteins can be used to generate antibodies
XX which are used to purify, detect and target the polypeptides, including
XX both in vivo and in vitro diagnostic and therapeutic methods. The
XX proteins can be used to treat or prevent neural, immune system, muscular,
XX reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
XX proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
XX sequences used in the exemplification of the present invention.
XX
XX Sequence 68 AA:
XX
XX Query Match 56.4%; Score 44; DB 21; Length 68;
XX Best Local Similarity 61.5%; Pred. No. 12;
XX Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX
XX Oy 1 AAGEGPALYEDPP 13
XX | | | | |
XX | | | | |
XX Db 39 AAGQEPXCLEDPP 51
XX
XX RESULT 10
XX ID AAU47601
XX AAU47601 standard; Protein: 241 AA.
XX
XX AAU47601;

```

XX AC AA047601;  
XX DT 27-FEB-2002 (first entry)  
XX DE Propionibacterium acnes immunogenic protein #8497.  
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
XX KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
XX KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
XX KW dermatological; osteopathic; neuroprotectant.  
XX OS  
XX PN WO200181581-A2.  
XX PD 01-NOV-2001.  
XX PF 20-APR-2001; 2001WO-US12865.  
XX PR 21-APR-2000; 2000US-199047P.  
XX PR 02-JUN-2000; 2000US-208841P.  
XX PR 07-JUL-2000; 2000US-216747P.  
XX PA (CORI-) CORIXA CORP.  
XX PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX DR WPI; 2001-616774/71.  
XX DR N-PSDB; AAS59539.  
XX PT Propionibacterium acnes polypeptides and nucleic acids useful for  
XX PT vaccinating against and diagnosing infections, especially useful for  
XX PT treating acne vulgaris -  
XX PS  
XX PS Example 1; SEQ ID No 8796; 10699p; English.  
XX CC Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic  
XX CC polypeptides. The proteins and their associated DNA sequences are used in  
XX CC the treatment, prevention and diagnosis of medical conditions caused by  
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
XX CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
XX CC P. acnes is also involved in infections of bone, joints and the central  
XX CC nervous system, however it is particularly involved in the inflammatory  
XX CC lesions associated with acne vulgaris. A method for detecting the  
XX CC presence or absence of P. acnes in a patient comprises contacting a  
XX CC sample with a binding agent that binds to the proteins of the invention  
XX CC and determining the amount of bound protein in the sample. The  
XX CC polypeptides may be used as antigens in the production of antibodies  
XX CC specific for P. acnes proteins. These antibodies can be used to  
XX CC downregulate expression and activity of P. acnes polypeptides and  
XX CC therefore treat P. acnes infections. The antibodies may also be used as  
XX CC diagnostic agents for determining P. acnes presence. For example, by  
XX CC enzyme linked immunosorbent assay (ELISA).  
XX CC Note: The sequence data for this patent did not form part of the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 241 AA;  
XX  
XX Query Match 55.1%; Score 43; DB 22; Length 241;  
XX Best Local Similarity 80.0%; Pred. No. 62;  
XX Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
XX  
XX QY 3 GEGPALYEDP 12  
XX 11 111 111  
XX Db 104 GEAPALVEDP 113

RESULT 11  
AAG53685  
ID AAG53685 standard; Protein; 999 AA.

XX AC AAG53685;  
XX DT 18-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 68372.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX PR 25-FEB-1999; 99US-0121825.  
XX PR 05-MAR-1999; 99US-0123160.  
XX PR 09-MAR-1999; 99US-0123548.  
XX PR 23-MAR-1999; 99US-0125788.  
XX PR 25-MAR-1999; 99US-0126264.  
XX PR 29-MAR-1999; 99US-0126785.  
XX PR 01-APR-1999; 99US-0127462.  
XX PR 06-APR-1999; 99US-0128234.  
XX PR 08-APR-1999; 99US-0128714.  
XX PR 16-APR-1999; 99US-0129845.  
XX PR 19-APR-1999; 99US-0130077.  
XX PR 21-APR-1999; 99US-0130449.  
XX PR 23-APR-1999; 99US-0130510.  
XX PR 23-APR-1999; 99US-0130891.  
XX PR 28-APR-1999; 99US-0131449.  
XX PR 30-APR-1999; 99US-0132048.  
XX PR 30-APR-1999; 99US-0132407.  
XX PR 04-MAY-1999; 99US-0132484.  
XX PR 05-MAY-1999; 99US-0132485.  
XX PR 06-MAY-1999; 99US-0132486.  
XX PR 07-MAY-1999; 99US-0132487.  
XX PR 08-MAY-1999; 99US-0132863.  
XX PR 11-MAY-1999; 99US-0134256.  
XX PR 14-MAY-1999; 99US-0134218.  
XX PR 14-MAY-1999; 99US-0134219.  
XX PR 14-MAY-1999; 99US-0134221.  
XX PR 14-MAY-1999; 99US-0134370.  
XX PR 18-MAY-1999; 99US-0134768.  
XX PR 19-MAY-1999; 99US-0134941.  
XX PR 20-MAY-1999; 99US-0135124.  
XX PR 21-MAY-1999; 99US-0135353.  
XX PR 24-MAY-1999; 99US-0135629.  
XX PR 25-MAY-1999; 99US-0136021.  
XX PR 27-MAY-1999; 99US-0136392.  
XX PR 28-MAY-1999; 99US-0136782.  
XX PR 01-JUN-1999; 99US-0137222.  
XX PR 03-JUN-1999; 99US-0137528.  
XX PR 04-JUN-1999; 99US-0137502.  
XX PR 07-JUN-1999; 99US-0137724.  
XX PR 08-JUN-1999; 99US-0138094.  
XX PR 10-JUN-1999; 99US-0138540.  
XX PR 10-JUN-1999; 99US-0138847.  
XX PR 14-JUN-1999; 99US-0139119.  
XX PR 16-JUN-1999; 99US-0139452.  
XX PR 17-JUN-1999; 99US-0139453.  
XX PR 17-JUN-1999; 99US-0139492.  
XX PR 18-JUN-1999; 99US-0139454.  
XX PR 18-JUN-1999; 99US-0139455.  
XX PR 18-JUN-1999; 99US-0139456.  
XX PR 18-JUN-1999; 99US-0139457.  
XX PR 18-JUN-1999; 99US-0139458.  
XX PR 18-JUN-1999; 99US-0139459.  
XX PR 18-JUN-1999; 99US-0139460.  
XX PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139889.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140685.  
PR 24-JUN-1999; 99US-0140686.  
PR 28-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142035.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144684.  
PR 20-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148365.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.

PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151085.  
PR 27-AUG-1999; 99US-0151086.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161982.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 55.1% Score 43; DB 21; Length 999;  
Best Local Similarity 58.3% Pred. No. 2.6e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 3 GEGPALYEDPPD 14  
| | | | | | | | | |  
Db 356 GGPSPHLQLPD 367

RESULT 12  
AAG53684  
ID AAG53684 standard; Protein: 1008 AA.

XX AAG53684;

XX 18-OCT-2000 (first entry)

DE Arelidopsis thaliana protein fragment SEQ ID NO: 68371.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
XX EP1033405-A2.  
PN  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126254.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145313.  
PR 27-JUL-1999; 99US-0145318.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.

PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0156458.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158292.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159684.  
PR 18-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160880.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160981.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match Best Local Similarity 55.14; Score 43; DB 21; Length 1008;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 3 GEGPALYEDPPD 14  
DB 365 GPGPSLHQLPPD 376

RESULT 13  
AAG53683  
ID AAG53683 standard; Protein: 1065 AA.

AC AAG53683;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 68370.

KW Protein identification; signal transduction pathway; metabolic pathway;

KM hybridisation assay; genetic mapping; gene expression control; promoter;

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
PR 09-MAR-1999; 99US-0123180.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 21-APR-1999; 99US-0130077.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 07-MAY-1999; 99US-0132486.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 20-MAY-1999; 99US-0134941.  
PR 21-MAY-1999; 99US-0135124.  
PR 24-MAY-1999; 99US-0135353.  
PR 25-MAY-1999; 99US-0135629.  
PR 27-MAY-1999; 99US-0136021.  
PR 28-MAY-1999; 99US-0136392.  
PR 03-JUN-1999; 99US-0136782.  
PR 04-JUN-1999; 99US-0137528.  
PR 07-JUN-1999; 99US-0137502.  
PR 08-JUN-1999; 99US-0137724.  
PR 10-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138340.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139465.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 22-JUN-1999; 99US-0139817.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 28-JUN-1999; 99US-0140695.  
PR 29-JUN-1999; 99US-0140823.  
PR 30-JUN-1999; 99US-0140991.  
PR 01-JUL-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151348.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.

PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159325.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 55.1%; Score 43; DB 21; Length 1065;  
Best Local Similarity 58.3%; Pred. No. 2.7e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 GEGPALYEDPPD 14  
DB 422 GEGPSLHQPPD 433

RESULT 14  
ID AAY21361 standard; protein; 115 AA.  
XX AAY21361;

DT 22-JUL-1999 (first entry)  
XX

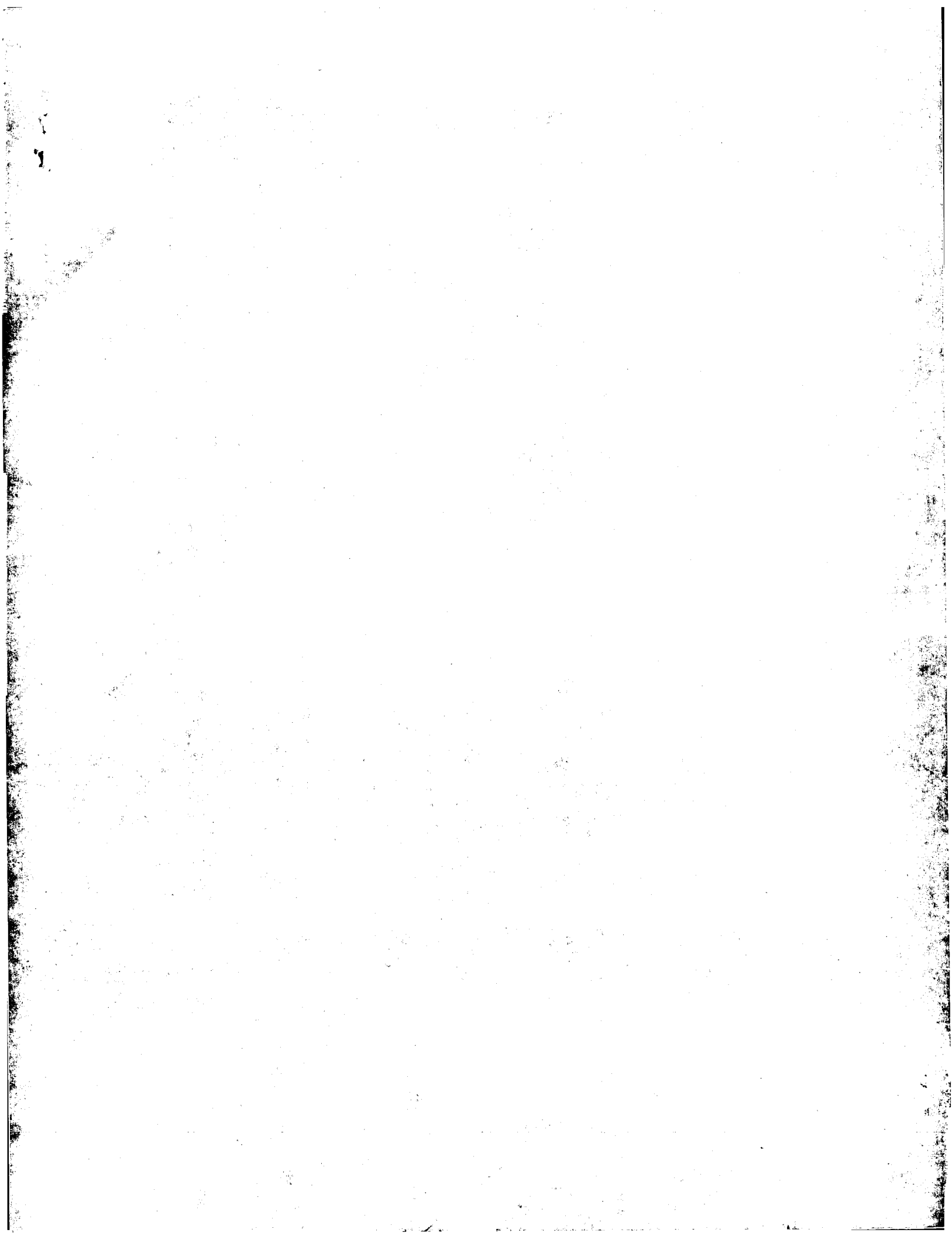
DE Human HUPF-I mutant protein fragment 13.

XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;  
KW frameshift mutation; age-related disease; neurodegenerative disorder;  
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;  
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;  
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;  
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;  
KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;  
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-I;  
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;  
XX high mobility group protein-C; neuroendocrine specific protein A.

OS Synthetic.  
OS Homo sapiens.  
XX

PN WO9845322-A2.  
 XX  
 PD 15-OC9-1998.  
 XX  
 PF 02-APR-1998; 98WO-1B00705.  
 XX  
 PR 10-APR-1997; 97US-0043163.  
 XX  
 PA (UYUT-) RIJXSUNIV UTRECHT  
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.  
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.  
 XX  
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;  
 XX  
 DR WPI: 1998-609901/51.  
 DR N-PSDB: AAX75768.  
 XX  
 PT Diagnosing disease by detecting frameshift mutations in RNA or  
 PT corresponding protein mutations - used to diagnose cancer and  
 PT neurological diseases, particularly Alzheimer's disease, and also  
 PT for treatment and prevention with specific ribozymes or wild-type  
 PT RNA  
 XX  
 PS Disclosure; Figure 17; 258pp; English.  
 XX  
 CC This invention describes a novel method for the diagnosis of a disease  
 CC caused by, or associated with, an RNA molecule that has a frameshift  
 CC mutation. The method is used to diagnose age-related diseases, especially  
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's  
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,  
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II  
 CC and many others listed) or susceptibility to these disorders. The method  
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,  
 CC at an early stage. It is based on the observation that disease may be  
 CC caused by mutations in RNA rather than DNA. The invention describes the  
 CC use of neuronal system RNA molecules, specifically proteins including the  
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated  
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule  
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,  
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic  
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma  
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group  
 CC protein-C (HMGp-C) and neuroendocrine specific protein A.  
 CC  
 XX  
 SQ Sequence 115 AA;  
 Query Match 53.8%; Score 42; DB 19; Length 115;  
 Best Local Similarity 80.0%; Pred. No. 42;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 OY 5 GPALYEDPPD 14  
 III IIIII  
 Db 94 GPANGEDPPD 103  
 RESULT 15  
 ABB97554  
 ID ABB97554 standard; Protein: 383 AA.  
 XX  
 AC ABB97554;  
 XX  
 DT 27-JUN-2002 (first entry)  
 XX  
 DE Novel human protein SEQ ID NO: 822.  
 XX  
 KW Human; anti-neurotic; anti-inflammatory; immunomodulator;  
 KW antifertility; cerebroprotective; cytoskeletal; rheumatic; gene therapy;  
 KW neuroprotective; antiparkinsonian; protein therapy; EST;  
 KW expressed sequence tag.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO980022660-A2.

XX  
 PD 21-MAR-2002.  
 XX  
 PF 10-SEP-2001; 2001WO-US26015.  
 XX  
 PR 11-SEP-2000; 2000US-0659671.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
 XX  
 DR WPI: 2002-292408/33.  
 DR N-PSDB: ABN32740.  
 XX  
 PT An isolated polynucleotide for treating diseases associated with its  
 PT encoded polypeptide such as cancer and multiple sclerosis -  
 XX  
 PS Claim 20; SEQ ID NO 822; 509pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of 444  
 CC novel human proteins. These were isolated from expressed sequences tags  
 CC (ESTs). They can be used to stimulate cell growth, to regulate  
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat  
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat  
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions  
 CC e.g. Rheumatoid arthritis, and to treat nervous system disorders e.g.  
 CC Parkinson's disease. The present sequence is a protein of the invention.  
 XX  
 SQ Sequence 383 AA;  
 Query Match 53.8%; Score 42; DB 23; Length 383;  
 Best Local Similarity 58.3%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 OY 1 AAGEGPALYEDP 12  
 :IIIIII:II  
 Db 164 SAGEGPVIVYHP 175  
 Search completed: January 30, 2003, 14:37:38  
 Job time : 37 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2003, 14:35:08 ; Search time 28 Seconds  
(without alignments)

103.024 Million cell updates/sec

Title: US-10-001-426-2

Perfect score: 78  
Sequence: 1 AAGEGSPALYEDPPD 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriopl:\*
- 17: sp\_archaeop:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	96.2	69	6	097870 ateles belz
2	74	94.9	289	11	099PF3 ratius norv
3	64	82.1	317	11	092232 cricetus
4	45	57.7	121	12	072205 hepatitis c
5	45	57.7	121	12	072206 hepatitis c
6	45	57.7	121	12	072207 hepatitis c
7	45	57.7	121	12	072208 hepatitis c
8	45	57.7	121	12	072209 hepatitis c
9	45	57.7	121	12	072210 hepatitis c
10	45	57.7	121	12	072211 hepatitis c
11	45	57.7	186	12	091X0 hepatitis c
12	45	57.7	186	12	091X6 hepatitis c
13	45	57.7	186	12	091X3 hepatitis c
14	45	57.7	186	12	091X2 hepatitis c
15	45	57.7	186	12	091X1 hepatitis c
16	45	57.7	186	12	091X9 hepatitis c

17	45	57.7	186	12	091X8 hepatitis c
18	45	57.7	186	12	091X4 hepatitis c
19	45	57.7	438	2	054823 streptomyce
20	44	56.4	328	16	08YC05 brucella me
21	43.5	55.8	612	16	09X010 thermotoga
22	43	55.1	121	12	09W682 hepatitis c
23	43	55.1	121	12	09W685 hepatitis c
24	43	55.1	121	12	09W687 hepatitis c
25	43	55.1	121	12	09W688 hepatitis c
26	43	55.1	186	12	091X0 hepatitis c
27	43	55.1	186	12	091X2 hepatitis c
28	43	55.1	186	12	091X8 hepatitis c
29	43	55.1	186	12	091X7 hepatitis c
30	43	55.1	186	12	091X6 hepatitis c
31	43	55.1	186	12	091X5 hepatitis c
32	43	55.1	224	12	073450 human papil
33	43	55.1	1006	10	08RX09 arabisdopsis
34	43	55.1	1157	17	08TK13 methanosarc
35	43	55.1	3104	2	004846 methanosarc
36	42.5	54.5	226	11	09DA92 mus musculu
37	42	53.8	157	6	095M7 equus caball
38	42	53.8	186	12	091X7 hepatitis c
39	42	53.8	186	12	091X05 hepatitis c
40	42	53.8	303	4	09UBES homo sapien
41	42	53.8	331	16	0920A9 rhlzobium m
42	42	53.8	392	11	0921Z6 rattus norv
43	42	53.8	583	5	09YH5 drosophila
44	42	53.8	945	6	09NI75 ovis aries
45	42	53.8	1017	4	09H9B8 homo sapien

# ALIGNMENTS

RESULT 1

ID 097870 PRELIMINARY: PRT: 69 AA.

AC 097870: 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE APEX nuclease (Fragment).

GN APX.

OS Ateles belzebuth chamek (Chamek spider monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateleinae; Ateles.

OX NCBI\_TaxID=118643;

RN [1]

RP SEQUENCE FROM N.A.

RA Canavez F.C., Senanez H.N.;

RT "Gene assignment in the spider monkey (Ateles paniscus chamek-APC):

RT APE-MHY7 to 2q; AR-GUA-F8C to the X chromosome.";

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF064685; AAD17542.1; -

DR HSSP: P27695; 1BIX.

DR InterPro: IPR000097; Apendonclsel.

DR InterPro: IPR005135; Exo\_endo\_phos.

DR Pfam: PF03372; Exo\_endo\_phos; 1.

DR PROSITE: PS00726; AP\_NUCLEASE\_F1.1; 1.

FT NON\_TER

FT NON\_TER

SQ SEQUENCE 69 AA: 7583 MW: 2FF30737F1581BD4 CRC64:

Query Match 96.2% Score 75; DB 6; Length 69;

Best local similarity 92.9% Pred. No. 3; 1e-05;

Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAGEGSPALYEDPPD 14

DB 1 AAGEGSPALYEDPPD 14

RESULT 2

```

099PF3
ID 099PF3 PRELIMINARY; PRT; 289 AA.
AC 099PF3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Apex (Fragment).
GN Apex.
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NORWAY;
RA Xie Z.H., Liu C.Z., He Y.H., Wang A.M., Ma C.;
RL Submitted (OCF-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF311054; AAC49922.1; -.
DR HSSP: P27695; 1BIX.
DR InterPro: IPR000097; Apendonclisel.
DR InterPro: IPR004442; ExoDNase_III.
DR InterPro: IPR004808; ExoIII_xth.
DR InterPro: IPR005135; Exo_endo_phos.
DR Pfam: PF03372; Exo_endo_phos; 1.
DR TIGRFAMS: TIGR00195; exoDNase_III; 1.
DR PROSITE: PS00726; AP_NUCLEASE_FL_1; 1.
DR PROSITE: PS00727; AP_NUCLEASE_FL_2; 1.
DR PROSITE: PS00728; AP_NUCLEASE_FL_3; 1.
FT NON-TER 1
SQ SEQUENCE 289 AA; 32353 MW; 67E82454D062CE51 CRC64;

Query Match 94.9%; Score 74; DB 11; Length 289;
Best Local Similarity 92.9%; Pred. No. 0.00021;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAGEGPALYEDPPD 14
DB 31 AAGEGPALYEDPPD 44

RESULT 3
Q92ZJ2 PRELIMINARY; PRT; 317 AA.
AC 092ZJ2;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Apurinic/apurimidinic endonuclease.
GN APE.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RX MEDLINE=20075118; PubMed=10606812;
RA Purohit S., Arenaz P.;
RT "Molecular cloning, sequence and structure analysis of hamster
apurinic/apurimidinic endonuclease (chapel) gene.";
RL Mutat. Res. 435:215-224(1999).
DR EMBL: AF056934; AAD0702.1; -.
DR HSSP: P27695; 1E9N.
DR InterPro: IPR000097; Apendonclisel.
DR InterPro: IPR004442; ExoDNase_III.
DR InterPro: IPR004808; ExoIII_xth.
DR InterPro: IPR005135; Exo_endo_phos.
DR Pfam: PF03372; Exo_endo_phos; 1.
DR TIGRFAMS: TIGR00195; exoDNase_III; 1.
DR PROSITE: PS00726; AP_NUCLEASE_FL_1; 1.
DR PROSITE: PS00727; AP_NUCLEASE_FL_1; 1.

```

```

DR PROSITE: PS00727; AP_NUCLEASE_FL_2; 1.
DR PROSITE: PS00728; AP_NUCLEASE_FL_3; 1.
KW Endonuclease.
SQ SEQUENCE 317 AA; 35512 MW; 76DD6F3975455408 CRC64;

Query Match 82.1%; Score 64; DB 11; Length 317;
Best Local Similarity 85.7%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAGEGPALYEDPPD 14
DB 36 AAGEGPALYEDAPD 49

RESULT 4
Q72205 PRELIMINARY; PRT; 121 AA.
ID 072205;
AC 072205;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ANTI-D RECIPIENT 1;
RX MEDLINE=98241727; PubMed=9573256;
RA McAllister J., Casino C., Davidson F., Power J., Lawlor E., Yap P.L.,
RT Simmonds P., Smith D.B.;
RT "Long-term evolution of the hypervariable region of hepatitis C virus
in a common-source-infected cohort.";
RL J. Virol. 72:4893-4905(1998).
CC -I- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR EMBL: AF056772; AAC17268.1; -.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NS1.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR ProDom: PD186062; HCV_NS1; 1.
DR Coat protein; Envelope protein; glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON-TER 1
FT NON-TER 121
SQ SEQUENCE 121 AA; 13281 MW; 8401C2D7374FDC8D CRC64;

Query Match 57.7%; Score 45; DB 12; Length 121;
Best Local Similarity 57.1%; Pred. No. 5.9;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAGEGPALYEDPPD 14
DB 101 AOGWGPITYAEPD 114

RESULT 5
Q72206 PRELIMINARY; PRT; 121 AA.
ID 072206;
AC 072206;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN-ANTI-D RECIPIENT 1;  
RX MEDLINE-98241727; PubMed-9573256;  
RA McAllister J., Casino C., Davidson F., Power J., Lawlor E., Yap P.L.,  
RT Simmonds P., Smith D.B.;  
RT "Long-term evolution of the hypervariable region of hepatitis C virus  
in a common-source-infected cohort.";  
RL J. Virol. 72:4893-4905(1998).  
CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.  
DR EMBL: AF056773; AAC17269.1; -.  
DR InterPro: IPR002519; HCV\_env.  
DR InterPro: IPR002531; HCV\_NSI.  
DR Pfam: PF01539; HCV\_env; 1.  
DR Pfam: PF01560; HCV\_NSI; 1.  
DR ProDom: PD186062; HCV\_NSI; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KM Polypeptide; Transmembrane.  
FT NON\_TER 1  
FT SEQUENCE 121 AA; 13281 MW; 8401C2D7374FDC8D CRC64;

Query Match 57.7%; Score 45; DB 12; Length 121;  
Best Local Similarity 57.1%; Pred. No. 5.9;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 AAGEGALYEDPPD 14  
| | | | | : | | |  
Db 101 AAGMGPTTVAEPPD 114

## RESULT 6

072207 PRELIMINARY; PRT; 121 AA.  
AC 072207;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Genome polyprotein [contains: envelope glycoprotein E2 (GP68) (GP70)  
DE (NS1)] (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage: Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ANTI-D RECIPIENT 1;  
RX MEDLINE-98241727; PubMed-9573256;  
RA McAllister J., Casino C., Davidson F., Power J., Lawlor E., Yap P.L.,  
RT Simmonds P., Smith D.B.;  
RT "Long-term evolution of the hypervariable region of hepatitis C virus  
in a common-source-infected cohort.";  
RL J. Virol. 72:4893-4905(1998).  
CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.  
DR EMBL: AF056774; AAC17270.1; -.  
DR InterPro: IPR002519; HCV\_env.  
DR InterPro: IPR002531; HCV\_NSI.  
DR Pfam: PF01539; HCV\_env; 1.  
DR Pfam: PF01560; HCV\_NSI; 1.  
DR ProDom: PD186062; HCV\_NSI; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KM Polypeptide; Transmembrane.  
FT NON\_TER 1  
FT SEQUENCE 121 AA; 13254 MW; B04523A965CA1367 CRC64;

Query Match 57.7%; Score 45; DB 12; Length 121;  
Best Local Similarity 57.1%; Pred. No. 5.9;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 AAGEGALYEDPPD 14  
| | | | | : | | |  
Db 101 AAGMGPTTVAEPPD 114

## RESULT 7

072208 PRELIMINARY; PRT; 121 AA.  
AC 072208;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Genome polyprotein [contains: envelope glycoprotein E2 (GP68) (GP70)  
DE (NS1)] (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage: Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ANTI-D RECIPIENT 1;  
RX MEDLINE-98241727; PubMed-9573256;  
RA McAllister J., Casino C., Davidson F., Power J., Lawlor E., Yap P.L.,  
RT Simmonds P., Smith D.B.;  
RT "Long-term evolution of the hypervariable region of hepatitis C virus  
in a common-source-infected cohort.";  
RL J. Virol. 72:4893-4905(1998).  
CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.  
DR EMBL: AF056775; AAC17271.1; -.  
DR InterPro: IPR002519; HCV\_env.  
DR InterPro: IPR002531; HCV\_NSI.  
DR Pfam: PF01539; HCV\_env; 1.  
DR Pfam: PF01560; HCV\_NSI; 1.  
DR ProDom: PD186062; HCV\_NSI; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KM Polypeptide; Transmembrane.  
FT NON\_TER 1  
FT SEQUENCE 121 AA; 13284 MW; A2C7B3A0FC5A087C CRC64;

Query Match 57.7%; Score 45; DB 12; Length 121;  
Best Local Similarity 57.1%; Pred. No. 5.9;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 AAGEGALYEDPPD 14  
| | | | | : | | |  
Db 101 AAGMGPTTVAEPPD 114

## RESULT 8

072209 PRELIMINARY; PRT; 121 AA.  
AC 072209;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Genome polyprotein [contains: envelope glycoprotein E2 (GP68) (GP70)  
DE (NS1)] (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage: Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ANTI-D RECIPIENT 1;  
RX MEDLINE-98241727; PubMed-9573256;  
RA McAllister J., Casino C., Davidson F., Power J., Lawlor E., Yap P.L.,  
RT Simmonds P., Smith D.B.;  
RT "Long-term evolution of the hypervariable region of hepatitis C virus  
in a common-source-infected cohort.";  
RL J. Virol. 72:4893-4905(1998).  
CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.  
DR EMBL: AF056776; AAC17272.1; -.  
DR InterPro: IPR002519; HCV\_env.  
DR InterPro: IPR002531; HCV\_NSI.  
DR Pfam: PF01539; HCV\_env; 1.  
DR Pfam: PF01560; HCV\_NSI; 1.  
DR ProDom: PD186062; HCV\_NSI; 1.

KW	Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KM	Polyprotein; Transmembrane.
FT	NON_TER 1
FT	NON_TER 121
SO	SEQUENCE 121 AA; 13268 MW; A2C7B3BBE75A1367 CRC64;
 Query Match	
Best Local Similarity	57.7%; Score 45; DB 12; Length 121;
Matches 8; Conservative	1; Mismatches 5; Indels 0; Gaps 0;
OY	1 AAGEGPALYEPPD 14         :
Db	101 AQGMCPTTYAEPPD 114
 RESULT 9	
O72210	PRELIMINARY; PRT; 121 AA.
ID	072210
AC	072210:
DT	01-AUG-1998 (TREMBLrel. 07, Created)
DT	01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE	Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment).
OS	Hepatitis C virus.
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
CC	Hepadnavirus.
OX	NCBI_TaxID=11103;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-ANTI-D RECIPIENT 1;
RA	MEDLINE=98241727; PubMed=9573256;
RX	McAllister J., Casino C., Davidson F., Power J., Lawlor E., Yap P.L., Simmonds P., Smith D.B.;
RT	"long-term evolution of the hypervariable region of hepatitis C virus in a common-source-infected cohort.";
RL	J. Virol. 72:4893-4905(1998).
CC	-I- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR	EMBL; AF056777; AAC17273.1; -
DR	InterPro; IPR002519; HCV_env.
DR	InterPro; IPR002531; HCV_NS1.
DR	Pfam; PF01539; HCV_env; 1.
DR	Pfam; PF01560; HCV_NS1; 1.
DR	PRODOM; PD186062; HCV_NS1; 1.
KW	Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KM	Polyprotein; Transmembrane.
FT	NON_TER 1
FT	NON_TER 121
SO	SEQUENCE 121 AA; 13268 MW; A2C7B3BBE75A1367 CRC64;
 Query Match	
Best Local Similarity	57.7%; Score 45; DB 12; Length 121;
Matches 8; Conservative	1; Mismatches 5; Indels 0; Gaps 0;
OY	1 AAGEGPALYEPPD 14         :
Db	101 AQGMCPTTYAEPPD 114
 RESULT 10	
O72211	PRELIMINARY; PRT; 121 AA.
ID	072211
AC	072211:
DT	01-AUG-1998 (TREMBLrel. 07, Created)
DT	01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE	Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment).
OS	Hepatitis C virus.
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
CC	Hepadnavirus.
OX	NCBI_TaxID=11103;
RN	[1]

RP	SEQUENCE FROM N.A.
RC	STRAIN=ANTI-D RECIPIENT 1;
RX	MEDLINE=98241727; PubMed=9573256;
RA	McAllister J., Casino C., Davidson F., Power J., Lawlor E., Yap P.L.,
RA	Simmonds P., Smith D.B.;
RT	"Long-term evolution of the hypervariable region of hepatitis C virus
rt	in a common-source-infected cohort.";
RL	J. Virol. 72:4893-4905(1998).
CC	-I- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR	EMBL; AF056778; AAC17274.1; -
DR	InterPro; IPR002519; HCV_env.
DR	InterPro; IPR002531; HCV_NSI.
DR	Pfam; PF01539; HCV_env. 1.
DR	Pfam; PF01560; HCV_NSI. 1.
DR	Prodrom; PD186062; HCV_NSI. 1.
KW	Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KM	Polyprotein; Transmembrane.
FT	NON_TER
FT	NON_TER
SO	SEQUENCE 121 AA; 13281 MW; 8401C2D737AFDC8D CRC64;
QY	Query Match 57.7%; Score 45; DB 12; Length 121; Best Local Similarity 57.1%; Pred. No. 5.9; Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
DB	1 AAGEGPALYEDPPD 14         :     101 AQGWPITYAEPPD 114
RESULT 11	
ID	G91XXO PRELIMINARY; PRT; 186 AA.
AC	G91XXO;
DT	01-OCT-2000 (TREMBLrel. 15, Created)
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE	Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment).
DE	Hepatitis C virus.
OS	Hepatitis C virus.
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC	Hepadnavirus.
OX	NCBI_Taxid-11103;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=P10;
RX	MEDLINE=20230065; PubMed=10764648;
RA	Farci P., Shioda A., Colana A., Diaz G., Peddis G., Meipolder J.C.,
RA	Strazzeria A., Chlen D.Y., Munoz S.J., Balestrieri A., Purcell R.H.,
RA	Alter H.J.;
RT	"The outcome of acute hepatitis C predicted by the evolution of the
RT	viral quasispecies.";
RL	Science 288:339-344(2000).
CC	-I- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR	EMBL; AF246078; AAF66377.1; -
DR	InterPro; IPR002519; HCV_env.
DR	InterPro; IPR002531; HCV_NSI.
DR	Pfam; PF01539; HCV_env. 1.
DR	Pfam; PF01560; HCV_NSI. 1.
DR	Prodrom; PD186062; HCV_NSI. 1.
KW	Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KM	Polyprotein; Transmembrane.
FT	NON_TER
FT	NON_TER
SO	SEQUENCE 186 AA; 20273 MW; DD4E1A0B4062A7BE CRC64;
QY	Query Match 57.7%; Score 45; DB 12; Length 186; Best Local Similarity 57.1%; Pred. No. 9.5; Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
DB	1 AAGEGPALYEDPPD 14         :     149 AQGWPITYAEPPD 162

## RESULT 12

091XW6 PRELIMINARY: PRT: 186 AA.  
 AC 091XW6;  
 DT 01-OCT-2000 (TEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
 DE 01-DEC-2001 (TEMBLrel. 19, Last annotation update)  
 DE Genome polyprotein [contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment).  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-P10;  
 RX MEDLINE-20230065; PubMed-10764648;  
 RA Farci P., Shimoda A., Colana A., Diaz G., Peddis G., Melpolder J.C., Strazzer A., Chien D.Y., Munoz S.J., Balestrieri A., Purcell R.H., Alter H.J.;  
 RT "The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies.";  
 RL Science 288:339-344(2000).  
 CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.  
 DR EMBL: AF246082; AAF66381.1; -;  
 DR InterPro: IPR002519; HCV\_env.  
 DR Pfam: PF01539; HCV\_env.1.  
 DR Pfam: PF01560; HCV\_NSI.1.  
 DR ProDom: PD186062; HCV\_NSI.1.  
 KM Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 FT NON\_TER 1  
 FT NON\_TER 186  
 SO SEQUENCE 186 AA; 20282 MW; 783A7DFCE1B41094 CRC64;

## Query Match

Best Local Similarity 57.7%; Score 45; DB 12; Length 186;  
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGECPALYEDPPD 14  
 DB 149 AOGMGPTVAEPD 162

## RESULT 13

091XW3 PRELIMINARY: PRT: 186 AA.  
 AC 091XW3;  
 DT 01-OCT-2000 (TEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
 DE 01-DEC-2001 (TEMBLrel. 19, Last annotation update)  
 DE Genome polyprotein [contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment).  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-P10;  
 RX MEDLINE-20230065; PubMed-10764648;  
 RA Farci P., Shimoda A., Colana A., Diaz G., Peddis G., Melpolder J.C., Strazzer A., Chien D.Y., Munoz S.J., Balestrieri A., Purcell R.H., Alter H.J.;  
 RT "The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies.";  
 RL Science 288:339-344(2000).  
 CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.  
 DR EMBL: AF246085; AAF66384.1; -;  
 DR InterPro: IPR002519; HCV\_env.

DR InterPro: IPR002531; HCV\_NSI.

DR Pfam: PF01539; HCV\_env.1.

DR Pfam: PF01560; HCV\_NSI.1.

DR ProDom: PD186062; HCV\_NSI.1.

KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;

KM Polyprotein; Transmembrane.

FT NON\_TER 1

FT NON\_TER 186

SO SEQUENCE 186 AA; 20280 MW; F056A403DEDFB1C CRC64;

## Query Match

Best Local Similarity 57.7%; Score 45; DB 12; Length 186;  
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGECPALYEDPPD 14  
 DB 149 AOGMGPTVAEPD 162

## RESULT 14

091XW2 PRELIMINARY: PRT: 186 AA.  
 AC 091XW2;  
 DT 01-OCT-2000 (TEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
 DE 01-DEC-2001 (TEMBLrel. 19, Last annotation update)  
 DE Genome polyprotein [contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment).  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-P10;  
 RX MEDLINE-20230065; PubMed-10764648;  
 RA Farci P., Shimoda A., Colana A., Diaz G., Peddis G., Melpolder J.C., Strazzer A., Chien D.Y., Munoz S.J., Balestrieri A., Purcell R.H., Alter H.J.;  
 RT "The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies.";  
 RL Science 288:339-344(2000).  
 CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.  
 DR EMBL: AF246086; AAF66385.1; -;  
 DR InterPro: IPR002519; HCV\_env.  
 DR Pfam: PF01539; HCV\_env.1.  
 DR Pfam: PF01560; HCV\_NSI.1.  
 DR ProDom: PD186062; HCV\_NSI.1.  
 KM Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 KW Polyprotein; Transmembrane.  
 FT NON\_TER 1  
 FT NON\_TER 186  
 SO SEQUENCE 186 AA; 20237 MW; CEC976E5527C741E CRC64;

## Query Match

Best Local Similarity 57.7%; Score 45; DB 12; Length 186;  
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGECPALYEDPPD 14  
 DB 149 AOGMGPTVAEPD 162

## RESULT 15

091XW1 PRELIMINARY: PRT: 186 AA.  
 AC 091XW1;  
 DT 01-OCT-2000 (TEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
 DE 01-DEC-2001 (TEMBLrel. 19, Last annotation update)  
 DE Genome polyprotein [contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment).

```

OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-P10;
RX MEDLINE=20230065; PubMed=10764648;
RA Farci P., Shimoda A., Coiana A., Diaz G., Peddis G., Melpolder J.C.,
RA Strazzer A., Chien D.Y., Munoz S.J., Balestrieri A., Purcell R.H.,
RA Alter H.J.;
RT "The outcome of acute hepatitis C predicted by the evolution of the
RT viral quasispecies.";
RT Science 288:339-344(2000).
CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR EMBL; AF246087; AAF66386.1; -.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NSL.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NSL; 1.
DR ProDom; PD186062; HCV_NSL; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 186
SQ SEQUENCE 186 AA; 20293 MW; B4D07F7DEF4EB197 CRC64;

Query Match 57.7%; Score 45; DB 12; Length 186;
Best Local Similarity 57.1%; Pred. No. 9.5;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAGGPGALYEDPPD 14
| | | | | : | | |
Db 149 AAGGPGITVAEPPD 162

```

Search completed: January 30, 2003, 14:38:31  
 Job time : 29 secs